

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 12:11:47 ; Search time 11 Seconds
(without alignments)
23.668 Million cell updates/sec

Title: US-09-871-974-2

Perfect score: 29

Sequence: 1 TKPPR 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	29	100.0	99	D127 HUMAN	Q911m4 homo sapien
2	29	100.0	201	COAE_SVNV3	Q55115 synchocyst
3	29	100.0	227	FA3C_HUMAN	Q92520 homo sapien
4	29	100.0	227	FA3C_MOUSE	Q91vri0 mus musculus
5	29	100.0	262	RPC_EPI163	P15238 bacterioph
6	29	100.0	465	ENGA_STRCO	Q9aww8 streptomyc
7	29	100.0	491	ENGA_STRAW	Q828y7 streptomyc
8	29	100.0	493	YE14 MYCPN	P75372 mycoplasma
9	29	100.0	604	YJ13 YEAST	P47030 saccharomyc
10	29	100.0	612	ELP1_MOUSE	Q60775 mus musculus
11	29	100.0	613	ELF1_HUMAN	P32519 homo sapien
12	29	100.0	919	DNLI_HUMAN	P18858 homo sapien
13	29	100.0	985	SECB_DROME	Q9vnh6 drosophila
14	29	100.0	1067	BAB2_DROME	Q9w0k4 drosophila
15	29	100.0	1235	TRK1 YEAST	P12685 saccharomyc
16	29	100.0	1241	TRK1_SACBA	P28569 saccharomyc
17	26	89.7	54	RS14_PYRAE	Q8xvml pyrobaculum
18	26	89.7	133	SECB_MOUSE	Q08535 mus musculus
19	26	89.7	133	VEGH_ORFN2	F52584 orf virus
20	26	89.7	140	RT12_DROME	P10735 drosophila
21	26	89.7	191	KGUA_SYNEL	Q8dmq7 synchococc
22	26	89.7	193	VCO7_ADE04	Q96831 human adeno
23	26	89.7	198	VCO7_ADE02	P03266 human adeno
24	26	89.7	199	KGUA_ANASP	Q8z017 anabaena sp
25	26	89.7	205	KGUA_NEIMA	Q9jtc96 neisseria m
26	26	89.7	205	KGUA_NEIME	Q9jtc95 neisseria m
27	26	89.7	219	Y132 NPVAC	P24730 autocograph
28	26	89.7	220	PE55_LUCUC	Q95ue8 lucilia cup
29	26	89.7	253	NOCT_RAT	Q9et55 rattus norv
30	26	89.7	257	TRUA_XYLFA	Q9pdke xyella fas
31	26	89.7	257	TRUA_XYLFA	Q87gs1 xyella fas
32	26	89.7	304	YQOB_CAEEL	Q09300 caenorhabdi
33	26	89.7	350	YB04_AQUAE	Q67189 aquifex aeo

34	26	89.7	370	TEMU_RICCN	Q92110 rickettsia
35	26	89.7	395	RT31_HUMAN	Q32665 homo sapien
36	26	89.7	396	TEMU_RHILO	Q38h10 rhizobium l
37	26	89.7	398	TEMU_BRUME	Q8y116 brucella me
38	26	89.7	398	TEMU_BRUSU	Q8y116 brucella su
39	26	89.7	402	VGLD_HSVEA	P24872 equine herp
40	26	89.7	428	CCKR_HUMAN	P32238 homo sapien
41	26	89.7	429	NOCT_MOUSE	Q35710 mus musculu
42	26	89.7	431	NOCT_HUMAN	Q9uk39 homo sapien
43	26	89.7	438	ENGA_CLOPE	Q8xjkl clostridium
44	26	89.7	439	SRMB_HAEIN	P44701 haemophilus
45	26	89.7	442	VGLD_HSVEB	P24844 equine herp
46	26	89.7	452	VGLD_HSVEB	P24379 equine herp
47	26	89.7	463	ENGA_BIFLO	Q8g6a8 bifidobacte
48	26	89.7	466	ENM5_HUMAN	Q00013 homo sapien
49	26	89.7	484	ENP6_HUMAN	Q75354 homo sapien
50	26	89.7	524	ANPA_AESPE	Q9y935 aeropyrium p
51	26	89.7	536	GAG_NLVB	P29168 murine leuk
52	26	89.7	552	AIRE_MOUSE	Q920e3 mus musculu
53	26	89.7	619	OM70_NEUCR	P23231 neurospora
54	26	89.7	737	POLG_HCVJ5	P27960 hepatitis c
55	26	89.7	737	POLG_HCVJ7	P27961 hepatitis c
56	26	89.7	782	CS2B_HUMAN	Q9h8e8 homo sapien
57	26	89.7	846	IRSL_HCVVA	P09715 human cytom
58	26	89.7	905	CIFF_MYCTU	Q10860 mycobacteri
59	26	89.7	928	CHS2_EXODE	P30601 exophialad
60	26	89.7	961	LIN2_CAEEL	P54936 caenorhabdi
61	26	89.7	961	TSP4_HUMAN	P35443 homo sapien
62	26	89.7	991	DHP1_SCHPO	P40848 schizosacch
63	26	89.7	1057	POLR_DROME	P16423 drosophila
64	26	89.7	1122	VAB1_CAEEL	O61460 caenorhabdi
65	26	89.7	1130	ITR6_HUMAN	P33229 homo sapien
66	26	89.7	1160	WDR7_HUMAN	Q9y4e6 homo sapien
67	26	89.7	1173	DPOL_RCMYM	Q85428 rat cytomeg
68	26	89.7	1446	I318_PRVFA	P33479 pseudorabie
69	26	89.7	1461	I318_PRVIF	P11675 pseudorabie
70	26	89.7	2139	CCAC_MOUSE	Q01815 mus musculu
71	26	89.7	2169	CCAC_RAT	P22002 rattus norv
72	26	89.7	2171	CCAC_RABIT	P15381 oryctolagus
73	26	89.7	2221	CCAC_HUMAN	P23936 homo sapien
74	26	89.7	2353	CCAH_HUMAN	Q95180 homo sapien
75	26	89.7	3033	POLG_HCVJ8	P24661 h genome po
76	26	89.7	5262	MLI2_HUMAN	O14686 homo sapien
77	25	86.2	44	COXR_RAT	P80433 rattus norv
78	25	86.2	69	COXR_BOVIN	P14622 bos taurus
79	25	86.2	111	LV21_HUMAN	P01712 homo sapien
80	25	86.2	114	MAUL_METEX	Q49129 methylobact
81	25	86.2	172	RUB2_PSEOL	P00272 pseudomonas
82	25	86.2	223	GLU2_MAIZE	P04706 zea mays (m
83	25	86.2	228	MX11_HUMAN	P50539 homo sapien
84	25	86.2	228	MX11_MOUSE	P50540 mus musculu
85	25	86.2	228	MX11_RAT	O09015 rattus norv
86	25	86.2	241	NGF_HUMAN	P01138 homo sapien
87	25	86.2	262	NEF_SIVM1	P05862 simian immu
88	25	86.2	277	TRMD_STRCO	O63882 streptomyc
89	25	86.2	278	PYRF_CORGL	Q8nq40 corynebacte
90	25	86.2	296	YWFH_BACSU	P39649 bacillus su
91	25	86.2	307	SGS4_DROME	Q00725 drosophila
92	25	86.2	307	FMT_BRUME	Q8ydb3 brucella me
93	25	86.2	311	FMT_AGR75	Q8uid0 agrobacteri
94	25	86.2	311	FMT_RHIME	Q92sh5 rhizobium m
95	25	86.2	317	FMT_RHILO	Q98d53 rhizobium l
96	25	86.2	318	IFR_CICAR	Q00016 ciccer ariet
97	25	86.2	332	SR4_PHYPO	P11113 physarum po
98	25	86.2	332	YIE2_HSVB4	Q02484 bovine herp
99	25	86.2	379	GI77_SCHPO	O59709 schizosacch
100	25	86.2	379	PSPB_DICDI	P54704 dictyosteli
101	25	86.2	396	TRT_DROME	P19351 drosophila
102	25	86.2	412	SU91_MOUSE	O54864 m histone-1
103	25	86.2	431	ACRO_RABIT	P48038 oryctolagus
104	25	86.2	434	UL43_HSV11	P10227 herpes simp
105	25	86.2	458	ENGA_HELPY	O25505 helicobacte
106	25	86.2	462	ENGA_HELPY	Q92109 helicobacte

107 25 86.2 469 1 ANTA_GENTR
108 25 86.2 487 1 CPDE_BOVIN
109 25 86.2 499 1 CPDE_CANFA
110 25 86.2 502 1 DCE_LYCES
111 25 86.2 609 1 YKD6_CABEL
112 25 86.2 647 1 TOP3_VIBCH
113 25 86.2 649 1 TOP3_SALTY
114 25 86.2 649 1 TOP3_SALTY
115 25 86.2 653 1 TOP3_ECOLI
116 25 86.2 672 1 NKR1_YEAST
117 25 86.2 679 1 ENV2_MOUSE
118 25 86.2 908 1 SRCA_RABIT
119 25 86.2 988 1 ST23_YEAST
120 25 86.2 1035 1 C68_YEAST
121 25 86.2 1074 1 EM12_MOUSE
122 25 86.2 1080 1 NKR1_YEAST
123 25 86.2 1152 1 ITAM_HUMAN
124 25 86.2 1175 1 OSRA_FAT
125 25 86.2 1210 1 AF4_HUMAN
126 25 86.2 1276 1 BMD_CLOBO
127 25 86.2 1411 1 TCOF_HUMAN
128 25 86.2 1484 1 CES2_HUMAN
129 25 86.2 1556 1 Y934_HUMAN
130 25 86.2 2146 1 INSR_DROME
131 25 86.2 2472 1 NCR2_MOUSE
132 25 86.2 2596 1 WDN1_HUMAN
133 24 82.8 57 1 Y23A_MYCLE
134 24 82.8 59 1 MAMB_DENJA
135 24 82.8 60 1 RZOD_ECOLI
136 24 82.8 61 1 TX51_DENJA
137 24 82.8 62 1 ITR1_ASCSU
138 24 82.8 63 1 YALA_ECOLI
139 24 82.8 67 1 SLVX_RHILLO
140 24 82.8 68 1 SLVX_BRUME
141 24 82.8 68 1 SLVX_BRUSU
142 24 82.8 73 1 UL12_HCMVA
143 24 82.8 80 1 SSS2_SCYCA
144 24 82.8 91 1 RL31_PYRAE
145 24 82.8 93 1 N12A_MEDSA
146 24 82.8 96 1 YBCO_ECOLI
147 24 82.8 103 1 N012_MEDTR
148 24 82.8 105 1 THGF_TOBAC
149 24 82.8 106 1 THG1_NICPA
150 24 82.8 107 1 REX1_SCHPO

Q9zw98 gentiana tr
Q01361 bos taurus
Q29473 canis famil
Q54767 lycopersico
Q03564 caenorhabdi
Q03565 vibrio chol
Q82655 salmonella
P40587 salmonella
P14294 escherichia
P32831 saccharomyc
P11370 mus musculu
P13666 oryctolagus
Q06010 saccharomyc
P32558 saccharomyc
Q84482 mus musculu
P36592 saccharomyc
P11315 homo sapien
P55266 rattus norv
P51825 homo sapien
P19321 clostridium
Q13428 homo sapien
Q9bxf3 homo sapien
Q9y2e4 homo sapien
P09208 drosophila
Q9w42 mus musculu
Q9w422 homo sapien
Q9cd20 mycobacteri
P28375 dendroaspis
P58041 escherichia
P01413 dendroaspis
P19398 ascaris suu
Q08366 escherichia
Q98gw1 rhizobium l
Q8yde6 brucella me
Q8tuz1 brucella su
P16777 human cytom
P11020 scyllorhinu
Q8ztu2 pyrobaculum
Q40361 medicago sa
Q47271 escherichia
P30365 medicago tr
P32026 nicotiana t
Q24115 nicotiana p
O13959 schizosacch

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P., Grahnam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E., Hammond S., Hartley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J., Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D., Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S., Lehmaelao M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D., Marsh V.L., Martin S.L., McConachie L.J., McMay K., McMurray A.A., Milne S.A., Mistry D., Moore W.J.P., Mullikin J.C., Nickerson T., Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I., Phillimore B.C.T., Prathalingam S.R., Plumb R.W., Ramsay H., Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S., Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E., Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A., Tracey A., Tronams A.C., Vaudin M., Wall M., Wallis J.M., Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A., Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S., Rogers J.;
The DNA sequence and comparative analysis of human chromosome 20.";
Nature 414:865-871 (2001).
[2]
SEQUENCE OF 19-77 FROM N.A., AND IDENTIFICATION.
TISSUE=B-cell, Fetal lung, and Testis;
MEDLINE=21843921; PubMed=11854508;
Schutte B.C., Mitros J.P., Bartlett J.A., Walters J.D., Jia H.P., Welsh M.J., Casavant T.L., McCray P.B. Jr.;
"Discovery of five conserved beta-defensin gene clusters using a computational search strategy.";
Proc. Natl. Acad. Sci. U.S.A. 99:2129-2133 (2002).
CC -!- FUNCTION: Has antibacterial activity (potential).
CC -!- SUBCELLULAR LOCATION: Secreted (potential).
CC -!- SIMILARITY: Belongs to the beta-defensin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL360078; CACU7685.1; -
CC EMBL; X122479; AAM3930.1; -
CC Genew; HGNC:16206; DEF127.
CC GO; GO:0003797; P:antibacterial peptide activity; TAS.
CC GO; GO:0045087; P:innate immune response; TAS.
CC -----
CC KW Antibiotic; Signal.
CC FT SIGNAL 1 18 POTENTIAL.
CC FT CHAIN 19 63 BETA-DEFENSIN 127.
CC FT PROPEP 66 99 POTENTIAL.
CC FT DISULFID 24 53 BY SIMILARITY.
CC FT DISULFID 33 47 BY SIMILARITY.
CC FT DISULFID 37 54 BY SIMILARITY.
CC FT CONFLICT 31 31 G -> R (IN REF. 2).
CC FT CONFLICT 71 77 RPKFATL -> SPKQHLH (IN REF. 2).
CC SQ SEQUENCE 99 AA; 11342 MW; C9CDB27D17AF380 CRC64;
Query Match 100.0%; Score 29; DB 1; Length 99;
Best Local Similarity 100.0%; Pred. No. 22; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;
QY 1 TKPPR 5
Db 67 TKPPR 71
RESULT 2
CORE_STY3
ID COAE SYNY3 STANDARD; PRT; 201 AA.
AC Q55515;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT Dephospho-CoA kinase (EC 2.7.1.24) (Dephosphocoenzyme A kinase).
DE

COAE OR SLR0553.
 Synechocystis sp. (strain PCC 6803).
 Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 NCBI_TaxID=1148;
 [1]
 SEQUENCE FROM N.A.
 MEDLINE=96127529; PubMed=8590279;
 Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
 Sugita M., Tabata S.,
 "Sequence analysis of the genome of the unicellular cyanobacterium
 Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
 region from map positions 64% to 92% of the genome."
 DNA Res. 2:153-166(1995).
 CC -1- FUNCTION: Catalyzes the phosphorylation of the 3'-hydroxyl group
 of dephosphocoenzyme A to form coenzyme A (By similarity).
 CC -1- CATALYTIC ACTIVITY: ATP + dephospho-CoA = ADP + CoA.
 CC -1- PATHWAY: Coenzyme A (CoA) biosynthesis; fifth (last) step.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the coae family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 EMBL; D64006; BAAL0873.1; -;
 HAMAP; MF 00376; -; 1;
 InterPro; IPR001977; Depp_CoKinase.
 Pfam; PF01121; Coae; 1. CoKinase; 1.
 ProDom; PD003329; Depp_CoKinase; 1.
 Trifam; TRF00152; TIGR00152; 1.
 PROSITE; PS01294; COAE; 1.
 KW Transferrase; Kinase; ATP-binding; Coenzyme A biosynthesis;
 Complete proteome.
 FT NP_BIND 15 22 ATP (POTENTIAL).
 SQ SEQUENCE 201 AA; 22520 MW; 34C3E142337F519E CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 201;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 |||||
 DB 3 TKPPR 7
 RESULT 3
 ID -FA3C HUMAN STANDARD; PRT; 227 AA.
 AC Q92520;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Protein FAM3C precursor (Protein GS3786).
 GN FAM3C.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone;
 AC Ono I., Hashimoto J., Takaoka K., Ochi T., Okubo K., Matsubara K.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RC Ryan E., Bauer C., Tucci S., Spalding L.;
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
 [3]

SEQUENCE FROM N.A.
 TISSUE=Ductum;
 MEDLINE=22388257; PubMed=12477932;
 Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Colling F.S., Wagner L., Sherman C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Busto K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalón D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Sahay J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 Whitney M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences."
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [4]
 CHARACTERIZATION.
 MEDLINE=22150867; PubMed=12160727;
 Zhu Y., Xu G., Patel A., McLaughlin M.M., Silverman C., Knecht K.A.,
 Sweitzer S., Li X., McDonnell P., Mirabile R., Zimmerman D., Boyce R.,
 Tierney L.A., Hu E., Livi G.P., Wolf B.A., Abdel-Meguid S.S.,
 Rose G.D., Aurora R., Hensley P., Briggs M., Young P.R.,
 "Cloning, expression, and initial characterization of a novel
 cytokine-like gene family."
 Genomics 80:144-150(2002).
 CC -1- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -1- TISSUE SPECIFICITY: Ubiquitous.
 CC -1- SIMILARITY: Belongs to the FAM3 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 EMBL; D87120; BAAL3251.1; -;
 EMBL; AC006364; AAQ96872.1; -;
 EMBL; BC046932; AAE46932.1; -;
 Genew; HGNC:18664; FAM3C.
 GO; GO:0005576; C:extracellular; NAS.
 GO; GO:0005125; F:cytokine activity; NAS.
 KW Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 227 PROTEIN FAM3C.
 FT DISULFID 58 221 POTENTIAL.
 FT DISULFID 64 86 POTENTIAL.
 SQ SEQUENCE 227 AA; 24680 MW; 6DC94B259052647F CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 227;
 Best Local Similarity 100.0%; Pred. No. 50;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 |||||
 DB 51 TKPPR 55
 RESULT 4
 ID -FA3C MOUSE STANDARD; PRT; 227 AA.
 AC Q91VU0; Q9CTB4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

28-FEB-2003 (Rel. 41, Created)
 28-FEB-2003 (Rel. 41, Last sequence update)
 15-MAR-2004 (Rel. 43, Last annotation update)
 GTP-binding protein engA.
 GN ENG_A OR SC01758 OR 28CI34.11C.
 Streptomyces coelicolor.
 OC Bacteria; Actinobacteridae; Actinomycetales;
 Streptomycineae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=1902;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 Seeger K., Saunders D., Sharp S., Squares S., Taylor K.,
 Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 Hopwood D.A.;
 "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2).";
 Nature 417:141-147(2002).
 RL -!- FUNCTION: GTPase of unknown physiological role.
 CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
 proteins. EngA subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AL939110; CAC12931.1; ALT_INIT.
 DR HAMAP; MF 00195; -; 1.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR005289; GTP-binding_dom.
 DR InterPro; IPR006073; GTP1_OBG.
 DR InterPro; IPR002917; MMR_HSR1.
 DR Pfam; PF01926; MMR_HSR1; 1.
 DR PRINTS; PR00326; GTP1_OBG.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00650; MG442; 2.
 DR TIGRFAMs; TIGR00231; small_GTP; 2.
 KW GTP-binding; Repeat; Complete proteome.
 FT NP_BIND 33 40 GTP 1 (POTENTIAL).
 FT NP_BIND 80 84 GTP 1 (POTENTIAL).
 FT NP_BIND 142 145 GTP 1 (POTENTIAL).
 FT NP_BIND 208 215 GTP 2 (POTENTIAL).
 FT NP_BIND 255 259 GTP 2 (POTENTIAL).
 FT NP_BIND 320 323 GTP 2 (POTENTIAL).
 FT NP_BIND 465 465 GTP 2 (POTENTIAL).
 SQ SEQUENCE 465 AA; 50344 MW; 95053E7C00C859F8 CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 465;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 414 TKPPR 418
 RESULT 7
 ENG_A_STRAW STANDARD; PRT; 491 AA.
 AC Q828Y7;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)

15-MAR-2004 (Rel. 43, Last annotation update)
 GTP-binding protein engA.
 GN ENG_A OR SAV6524.
 OC Streptomyces avermitilis.
 Bacteria; Actinobacteridae; Actinomycetales;
 Streptomycineae; Streptomycetaceae; Streptomyces.
 NCBI_TaxID=33903;
 [1]
 SEQUENCE FROM N.A.
 RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
 Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 "Genome sequence of an industrial microorganism Streptomyces
 avermitilis: deducing the ability of producing secondary
 metabolites.";
 Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 [2]
 SEQUENCE FROM N.A.
 RC STRAIN=NA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
 RX MEDLINE=22608306; PubMed=13692562;
 RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
 Sakaki Y., Hattori M., Omura S.;
 "Complete genome sequence and comparative analysis of the industrial
 microorganism Streptomyces avermitilis.";
 Nat. Biotechnol. 21:526-531(2003).
 RL -!- FUNCTION: GTPase of unknown physiological role.
 CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
 proteins. EngA subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP005047; BAC74235.1; -.
 DR HAMAP; MF 00195; -; 1.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR005289; GTP-binding_dom.
 DR InterPro; IPR006073; GTP1_OBG.
 DR InterPro; IPR002917; MMR_HSR1.
 DR Pfam; PF01926; MMR_HSR1; 1.
 DR PRINTS; PR00326; GTP1_OBG.
 DR SMART; SM00382; AAA; 2.
 DR TIGRFAMs; TIGR00650; MG442; 2.
 DR TIGRFAMs; TIGR00231; small_GTP; 2.
 KW GTP-binding; Repeat; Complete proteome.
 FT NP_BIND 59 66 GTP 1 (POTENTIAL).
 FT NP_BIND 106 110 GTP 1 (POTENTIAL).
 FT NP_BIND 168 171 GTP 1 (POTENTIAL).
 FT NP_BIND 234 241 GTP 2 (POTENTIAL).
 FT NP_BIND 281 285 GTP 2 (POTENTIAL).
 FT NP_BIND 346 349 GTP 2 (POTENTIAL).
 SQ SEQUENCE 491 AA; 53353 MW; 7BADDD8E0449D14AC CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 440 TKPPR 444
 RESULT 8
 YE14_MYCPN STANDARD; PRT; 493 AA.
 ID YE14_MYCPN
 AC P75372;

```
--!- SIMILARITY: TO YEAST YKR019C.
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; Z49358; CAA89376.1; -.
EMBL; X83502; CAA58487.1; -.
EMBL; X88851; CAA61316.1; -.
PIR; S56027; S56027.
GeneOnline; I41697; -.
SGD; S0003619; YJL083W.
InterPro; IPRO00261; EPSI5_homology.
SMART; SM00027; EH; 1.
Hypothetical protein.
SEQUENCE 604 AA; 68768 MW; EL7C6C71D9F8PBD6 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 604;
Best local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
||| |
DB 151 TKPPR 155

RESULT 10
ELF1_MOUSE
ID ELF1_MOUSE STANDARD; PRT; 612 AA.
AC Q60775;
DDT 15-JUN-1998 (Rel. 36, Created)
DI 15-JUL-1998 (Rel. 36, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE ETS-related transcription factor Elf-1 (E74-like factor 1).
DEF Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fibroblast;
RX MEDLINE=96257231; PubMed=8656284;
RA Davis J.N., Roussel M.F.;
RT "Cloning and expression of the murine Elf-1 cDNA."
RL Gene 171:265-269(1996).
CC --!
CC FUNCTION: TRANSCRIPTION FACTOR THAT APPEARS TO BE REQUIRED FOR THE
CC T-CELL-RECEPTOR-MEDIATED TRANS ACTIVATION OF HIV-2 GENE
CC EXPRESSION. BINDS SPECIFICALLY TO TWO PURINE-RICH MOTIFS IN THE
CC HIV-2 ENHANCER. ELF-1 BINDS TO THE UNDERPHOSPHORYLATED FORM OF RB.
CC MAY INTERACT WITH OTHER TRANSCRIPTION FACTORS IN ORDER TO REGULATE
CC SPECIFIC GENES (BY SIMILARITY).
CC --!
CC SUBCELLULAR LOCATION Nuclear
CC --!
CC TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN HEMATOPOIETIC CELLS.
CC DETECTED IN OTHER CELL TYPES SUCH AS FIBROBLASTS.
CC --!
CC SIMILARITY: Belongs to the ETS family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

EMBL; U19617; AAB17097.1; -.
HSSP; P28324; 1BC8.
DR TRANSFAC; T05012; -.
DR MGD; MG1:107180; Elf1.
DR

```

DR InterPro; IPR000418; ETS.
DR InterPro; IPR002341; HSF_ETs.
DR Pfam; PF00178; Ets; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS00361; ETS_DOMAIN_3; 1.
DR Nuclear protein; Transcription regulation; Activator; DNA-binding.
KW DOMAIN 75 80 POLY-ASP.
FT DNA_BIND 208 290 ETS-DOMAIN.
SQ SEQUENCE 612 AA; 66221 MW; 442F4C95142B31F0 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 612;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPFR 5
Db 180 TKPFR 184

RESULT 11
ELF1_HUMAN
ID ELF1_HUMAN STANDARD; PRT; 619 AA.
AC P32519; Q9UBEL;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE ETS-related transcription factor Elf-1 (E74-like factor 1).
GN ELF1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92407982; PubMed=1527846;
RA Leiden J.M., Wang C.Y., Petryniak B., Markovitz D.M., Nabel G.J.,
RA Thompson C.B.;
RA "A novel Ets-related transcription factor, Elf-1, binds to human
RT immunodeficiency virus type 2 regulatory elements that are required
RT for inducible trans activation in T cells.";
RL J. Virol. 66:5890-5897(1992).
RN [2]
RP SEQUENCE OF 204-289 FROM N.A.
RX MEDLINE=92186836; PubMed=1545787;
RA Thompson C.B., Wang C.Y., Ho I.C., Bohjanen P.R., Petryniak B.,
RA June C.H., Miesfeldt S., Zhang L., Nabel G.J., Karpinski B.;
RA "Cis-acting sequences required for inducible interleukin-2 enhancer
RT function bind a novel Ets-related protein, Elf-1.";
RL Mol. Cell. Biol. 12:1043-1053(1992).
RN [3]
RP BINDING TO RB.
RX MEDLINE=93262492; PubMed=8493578;
RA Wang C.Y., Petryniak B., Thompson C.B., Kaelin W.G., Leiden J.M.;
RA "Regulation of the Ets-related transcription factor Elf-1 by binding
RT to the retinoblastoma protein.";
RL Science 260:1330-1335(1993).
CC -!- FUNCTION: TRANSCRIPTION FACTOR THAT APPEARS TO BE REQUIRED FOR THE
CC T-CELL-RECEPTOR-MEDIATED TRANS ACTIVATION OF HIV-2 GENE
CC EXPRESSION. BINDS SPECIFICALLY TO TWO PURINE-RICH MOTIFS IN THE
CC HIV-2 ENHANCER. ELF-1 BINDS TO THE UNDERPHOSPHORYLATED FORM OF RB.
CC MAY INTERACT WITH OTHER TRANSCRIPTION FACTORS IN ORDER TO REGULATE
CC SPECIFIC GENES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the ETS family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----NOT_ANNOTATED_CDS-----
CC ENBL; M82882; -; NOT_ANNOTATED_CDS.
CC PIR; A43361; A43361.
DR HSP; P28324; LBC8.
DR TRANSFAC; T01113; -.
DR Genew; HGNC:3316; ELF1.
DR MTM; 189973; -.
DR GO; GO:0005634; C:nucleus; NAS.
DR GO; GO:0015653; F:transcriptional activator activity; NAS.
DR GO; GO:0008355; P:regulation of transcription, DNA-dependent; NAS.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETs.
DR Pfam; PF00178; Ets; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SM00413; ETS; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS00361; ETS_DOMAIN_3; 1.
KW Nuclear protein; Transcription regulation; Activator; DNA-binding.
FT DOMAIN 75 80 POLY-ASP.
FT DNA_BIND 208 290 ETS-DOMAIN.
FT CONFLICT 283 283 Q -> G (IN REF. 2).
SQ SEQUENCE 619 AA; 67455 MW; AB0B41B2964A66EF CRC64;

Query Match 100.0%; Score 29; DB 1; Length 619;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPFR 5
Db 180 TKPFR 184

RESULT 12
DNL1_HUMAN
ID DNL1_HUMAN STANDARD; PRT; 919 AA.
AC P18858;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DNA ligase I (EC 6.5.1.1) (Polydeoxyribonucleotide synthase [ATP]).
GN LIG1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=T lymphoblast;
RX MEDLINE=90370849; PubMed=2204063;
RA Barnes D.E., Johnston L.H., Kodama K.I., Tomkinson A.E.,
RA Lasko D.D., Lindahl T.;
RA "Human DNA ligase I cDNA: cloning and functional expression in
RT Saccharomyces cerevisiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:6679-6683(1990).
RN [2]
RP SEQUENCE FROM N.A.
RA Rieder M.J., Livingston R.J., Braun A.C., Montoya M.A., Chung M.-W.,
RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 716-753 FROM N.A.
RX MEDLINE=91352039; PubMed=1881902;
RA Petrini J.H.J., Huwiler K.G., Weaver D.T.;
RA "A wild-type DNA ligase I gene is expressed in Bloom's syndrome
RT cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:7615-7619(1991).
RN [4]
RP VARIANTS LYS-566 AND TRP-771.
RX MEDLINE=92257590; PubMed=1581963;

```

RA Barnes D.E., Tomkinson A.E., Lehmann A.R., Webster A.D.B.,
 RA Lindahl T.;
 RT "Mutations in the DNA ligase I gene of an individual with
 RT immunodeficiencies and cellular hypersensitivity to DNA-damaging
 RT agents.";
 RL Cell 69:495-503(1992).
 CC -!- FUNCTION: This protein seals, during DNA replication, DNA
 CC recombination and DNA repair, nicks in double-stranded DNA.
 CC -!- CATALYTIC ACTIVITY: ATP + (deoxyribonucleotide)(N) +
 CC (deoxyribonucleotide)(N-M)
 CC (deoxyribonucleotide)(N-M)
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- PM: Phosphorylated in vivo.
 CC -!- DISEASE: Defects in *lig1* seem to cause immunodeficiencies and
 CC cellular hypersensitivity to DNA-damaging agents.
 CC -!- SIMILARITY: Belongs to the ATP-dependent DNA ligase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M36067; AAA59518.1; -;
 DR EMBL: AF527418; AM77697.1; -;
 DR PIR: A36048; A41275.
 DR Genew: HGNC:6598; LIG1.
 DR MIM: 126391; -;
 DR GO: GO:0005634; C:nucleus; TAS.
 DR GO: GO:0003677; F:DNA binding; TAS.
 DR GO: GO:0003909; F:DNA ligase activity; TAS.
 DR GO: GO:0006281; P:DNA repair; TAS.
 DR GO: GO:0007345; P:embryogenesis and morphogenesis; TAS.
 DR InterPro: IPR000977; DNA_ligase.
 DR Pfam: PF01068; DNA_ligase_1.
 DR Pfam: PF04679; DNA_ligase_A_C; 1.
 DR Pfam: PF04675; DNA_ligase_A; 1.
 DR TIGRFAMs: TIGR00574; chl1; 1.
 DR PROSITE: PS00697; DNA_LIGASE_A1; 1.
 DR PROSITE: PS00333; DNA_LIGASE_A2; 1.
 DR PROSITE: PS0160; DNA_LIGASE_A3; 1.
 KW DNA repair; DNA replication; DNA recombination; Cell division; Ligase;
 KW ATP-binding; Nuclear protein; Phosphorylation; Polymorphism;
 KW Disease mutation.
 FT BINDING 568 AMP (BY SIMILARITY).
 FT VARIANT 249 G -> E (in dbSNP:3730911).
 FT VARIANT 267 /FTId=VAR_016766.
 FT VARIANT 267 N -> S (in dbSNP:3730933).
 FT VARIANT 409 /FTId=VAR_016767.
 FT VARIANT 409 R -> H (in dbSNP:4987068).
 FT VARIANT 480 /FTId=VAR_016768.
 FT VARIANT 480 M -> V (in dbSNP:3730980).
 FT VARIANT 566 /FTId=VAR_016769.
 FT VARIANT 566 E -> K (in *lig1* deficiency).
 FT VARIANT 614 T -> I (in dbSNP:3731003).
 FT VARIANT 771 R -> W (in *lig1* deficiency).
 FT VARIANT 771 /FTId=VAR_002263.
 SQ SEQUENCE 919 AA; 101735 MW; B2954D3E38A8D4D CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 919;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPR 5

Db 216 TKPR 222

RESULT 13

SECS_DROME
 ID SECS_DROME STANDARD; PRT; 985 AA.
 AC Q9VNH6; 2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable exocyst complex component Sec8.
 GN CG2095.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_taxid=7227;
 RN [1]
 SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA Amanatides P.G., Scherer S.E., Richards S., Aeburner M., Henderson S.N.,
 RA George R.A., Lewis S.E., Yandell M.D., Zhang Q., Chen L.Y.,
 RA Brandon R.C., Rogers Y.-H.C., Blazef R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter B.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agayani A., An H.-J., Andrews-Biankoc C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley S.M.,
 RA Beeson K.Y., Betos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Foslter C., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lako P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarly C., Morris J., Moshrefl A.,
 RA Mount S.M., Moy M., Murphy E., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J., Yao Q.A.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Zhao Q., Zheng L.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou X.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of *Drosophila melanogaster*.";
 Science 287:2185-2195(2000).
 [2]
 RN [2]
 REVISIONS.
 RP MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celinker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [3]
 SEQUENCE FROM N.A.
 RC STRAIN=Berkely; TISSUE=Embryo;
 RX MEDLINE=22426066; PubMed=12537569;
 RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champagne M.,

George R.A., Guarin H., Kronmiller B., Pacleab J.M., Park S., Wan K.H.,
 RA Rubin G.M., Celniker S.E.,
 RT "A Drosophila full-length cDNA resource";
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
 CC -!- FUNCTION: Component of the exocyst complex involved in the docking
 CC of exocystic vesicles with fusions site on the plasma membrane (By
 CC similarity).
 CC SUBUNIT: The exocyst complex is composed of SEC3, SEC5, SEC6,
 CC SEC8, SEC10, SEC15, EXO70 and EXO84 (By similarity).
 CC -!- SIMILARITY: Belongs to the SEC8 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB003601; AAF51959.3; -;
 CC EMBL; AY19660; AM50314.1; -;
 CC FlyBase; FBgn0037373; CG2095.
 CC InterPro; IPR004172; L27.
 CC InterPro; IPR007191; Sec8_exocyst.
 CC Pfam; PF02828; L27; 1.
 CC Pfam; PF04048; Sec8_exocyst; 1.
 CC DR Exocytosis; Transports; Protein transport; Coiled coil.
 CC KW DOMAIN 36 70
 CC CONFLICT 118 118 M -> T (IN REF. 3).
 CC CONFLICT 209 209 E -> D (IN REF. 3).
 CC CONFLICT 333 333 S -> T (IN REF. 3).
 CC CONFLICT 668 668 Y -> S (IN REF. 3).
 CC FT CONFLICT 668 668 Y -> S (IN REF. 3).
 CC SQ SEQUENCE 985 AA; 111665 MW; 6BFB0D9C539FBE7 CRC64;
 Query Match 100.0%; Score 29; DB 1; Length 985;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 7 TKPPR 11
 RESULT 14
 ID BAB2 DROME STANDARD; PRT; 1067 AA.
 AC Q9W0K4; Q24001; Q9UH3;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DE Bric-a-brac protein 2.
 GN BAB2 OR BTB1 OR CG9102/CG13911.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, AND TISSUE
 RP SPECIFICITY.
 RC TISSUE=Embryo, and Ovary;
 RX MEDLINE=21969340; PubMed=11973274;
 RA Couderc J.L.G., Godt D., Zolman S., Chen J., Li M., Tjong S.,
 RA Cranton S.E., Sanut-Barnola I., Laski F.A.;
 RT "The bric a brac locus consists of two paralogous genes encoding
 RT BTB/POZ domain proteins and acts as a homeotic and morphogenetic
 RT regulator of imaginal development in Drosophila";
 RL Development 129:2419-2433(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon K.C., Rogers J.H.C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.B., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fessler G., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Gicdek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpan G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP REVISIONS.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 RN [4]
 RP SEQUENCE OF 196-310 FROM N.A.
 RX MEDLINE=95024186; PubMed=7938017;
 RA Zolman S., Godt D., Prive G.G., Couderc J.L., Laski F.A.;
 RT "The BTB domain, found primarily in zinc finger proteins, defines an
 RT evolutionarily conserved family that includes several developmentally
 RT regulated genes in Drosophila";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:10717-10721(1994).
 CC -!- FUNCTION: Probably acts as a transcriptional regulator. Required
 CC for the specification of the tarsal segment. Also involved in
 CC antenna development.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -!- TISSUE SPECIFICITY: Leg imaginal disk at the central region of the
 CC tarsus and in eye antenna disk at the basal cylinder.
 CC -!- MISCELLANEOUS: 'bric-a-brac' means 'jumble' in French (referring to
 CC the mutant ovary phenotype).
 CC -!- SIMILARITY: Contains 1 A.T hook DNA-binding repeat.
 CC -!- SIMILARITY: Contains 1 BTB/POZ domain.
 CC -!- SIMILARITY: Contains 1 helix-turn-helix Pq-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AJ252173; CAB64388.1; -;
DR EMBL; AEO03470; AAF47442.2; -;
DR EMBL; U14399; AAS0834.1; -;
DR Flybase; FSN0025525; bab2.
DR InterPro; IPR000637; AT hook.
DR InterPro; IPR000210; BTB_POZ.
DR Pfam; PF00651; BTB; 1.
DR SMART; SM00225; BTB; 1.
DR PROSITE; PS00097; BTB; 1.
KW Nuclear protein; DNA-binding; Transcription regulation.
FT DOMAIN 223 288
FT DNA_BIND 645 690 H-T-H MOTIF PSQ-TYPE.
FT DNA_BIND 697 708 A-T HOOK.
FT CONFLICT 858 858 A -> R (IN REF. 1).
SQ SEQUENCE 1067 AA; 114661 MW; 7DBFC7681D507FC0 CRC64;

Query Match 100.0%; Score 29; DB 1; Length 1067;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TKPPR 5
Db 181 TKPPR 185

RESULT 15
TRK1 YEAST
ID -TRK1 YEAST STANDARD; PRT; 1235 AA.
AC P12685;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Potassium transport protein, high-affinity.
GN TRK1 OR YJL129C OR Y0693.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
RX MEDLINE=88302204; PubMed=3043197;
RA Gaber R.F., Styles C.A., Fink G.R.;
RT "TRK1 encodes a plasma membrane protein required for high-affinity
RT potassium transport in Saccharomyces cerevisiae.";
EL Mol. Cell. Biol. 8:2848-2859(1988).
RN [2]

RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FY1679;
RX MEDLINE=97103775; PubMed=8948101;
RA Cziepluch C., Kordes E., Fajol A., Jauniaux J.-C.;
RT "Sequencing analysis of a 40.2 kb fragment of yeast chromosome X
RT reveals 19 open reading frames including URA2 [5' end], TRK1, PBS2,
RT SP10, GDI4, RPE1, PHO86, NCA3, ASE1, COT7, GZF3, two tRNA genes,
RT three remnant delta elements and a Ty4 transposon.";
RL Yeast 12:1471-1474 (1996).
CC -1- FUNCTION: This protein is required for high-affinity potassium
CC transport.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the trkH potassium transport family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M21328; AAA34728.1; -;
DR EMBL; Z49404; CAA89424.1; -;
DR PIR; S05849; PWBHY.
DR Germline; 141741; -;
DR SGD; S0003665; TRK1.
DR GO; GO:0015079; P:potassium ion transporter activity; IDA.
DR GO; GO:0030007; P:potassium ion homeostasis; IDA.
DR InterPro; IPR003445; Cat_transp.
DR InterPro; IPR004773; Ktransp_euk.
DR Pfam; PF02386; TrkH; 1.
DR TIGRFAMs; TIGR002450; K+ transporter TRK; 1.
DR TIGRFAMs; TIGR00934; 2a38euk; 1.
KW Transport; Transmembrane; Potassium transport; Glycoprotein.
FT TRANSMEM 49 70 POTENTIAL.
FT TRANSMEM 78 98 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
FT TRANSMEM 778 800 POTENTIAL.
FT TRANSMEM 813 834 POTENTIAL.
FT TRANSMEM 838 859 POTENTIAL.
FT TRANSMEM 862 882 POTENTIAL.
FT TRANSMEM 898 918 POTENTIAL.
FT TRANSMEM 923 943 POTENTIAL.
FT TRANSMEM 971 991 POTENTIAL.
FT TRANSMEM 1078 1098 POTENTIAL.
FT TRANSMEM 1111 1131 POTENTIAL.
FT CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 169 169 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 222 222 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 369 369 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 383 383 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 501 501 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 677 677 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 919 919 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1135 1135 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1235 AA; 141072 MW; BCE9FD8A0BA0982B CRC64;
Query Match 100.0%; Score 29; DB 1; Length 1235;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TKPPR 5
Db 438 TKPPR 442

RESULT 16
TRK1 SACBA
ID -TRK1 SACBA STANDARD; PRT; 1241 AA.
AC P28569;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Potassium transport protein, high-affinity.
GN TRK1.
OS Saccharomyces bayanus (Yeast) (Saccharomycetes uvarum).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4931;
[1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=R1668;
RX MEDLINE=91216443; PubMed=2022322;
RA Anderson J.A., Best L.A., Gaber R.F.;
RT "Structural and functional conservation between the high-affinity K+
RT transporters of Saccharomyces uvarum and Saccharomyces cerevisiae.";
RL Gene 99:39-46 (1991).
CC -1- FUNCTION: This protein is required for high-affinity potassium

CC transport.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the trkH potassium transport family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M57508; AAA34661.1; -;
 CC InterPro; IPR003445; Cat. transp.
 CC InterPro; IPR004773; Ktransp_euk.
 CC Pfam; PF02386; TrkH; 1.
 CC PIRSF; PIRSF002450; K. transp. TRK; 1.
 CC TIGRFAMs; TIGR00934; 2a38euk; 1.
 CC TRANSFER; Transmembrane; Potassium transp.; Glycoprotein.
 CC TRANSMEM 49 70 POTENTIAL.
 CC TRANSMEM 78 98 POTENTIAL.
 CC TRANSMEM 107 127 POTENTIAL.
 CC TRANSMEM 784 806 POTENTIAL.
 CC TRANSMEM 819 840 POTENTIAL.
 CC TRANSMEM 844 864 POTENTIAL.
 CC TRANSMEM 868 888 POTENTIAL.
 CC TRANSMEM 904 924 POTENTIAL.
 CC TRANSMEM 929 949 POTENTIAL.
 CC TRANSMEM 977 997 POTENTIAL.
 CC TRANSMEM 1084 1104 POTENTIAL.
 CC TRANSMEM 1117 1137 POTENTIAL.
 CC CARBOHYD 100 100 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 223 223 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 227 227 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 233 233 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 257 257 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 274 274 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 353 353 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 389 389 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 442 442 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 505 505 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 538 538 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 584 584 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 660 660 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 681 681 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 691 691 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 741 741 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 925 925 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CARBOHYD 1141 1141 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC SEQUENCE 1241 AA; 141225 MW; F5149E47EC00BCD2 CRC64;
 CC -----
 CC Query Match 100.08; Score 29; DB 1; Length 1241;
 CC Best Local Similarity 100.08; Pred. No. 2.9e+02;
 CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC Qy 1 TKPPR 5
 CC |||||
 CC 444 TKPPR 448
 CC -----
 CC RESULT 17
 CC RS14 PYRAE STANDARD; PRT; 54 AA.
 CC AC Q8ZVM1; P58732;
 CC DT 28-FEB-2003 (Rel. 41, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE 30S ribosomal protein S14p.
 CC GN RPS14P OR PAE2097.
 CC OS Pyrobaculum aerophilum.
 CC OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 CC Thermoproteaceae; Pyrobaculum.

OX NCBI_TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX MEDLINES=21664397; PubMed=11792869;
 RA Fitz-Gibbon S.R., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
 aerophilum";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 CC -!- SIMILARITY: Belongs to the S14p family of ribosomal proteins.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AE009857; AAL63943.1; -;
 CC InterPro; IPR001209; Ribosomal_S14.
 CC Pfam; PF00253; Ribosomal_S14; 1.
 CC PROSITE; PS00527; RIBOSOMAL_S14; 1.
 CC Ribosomal protein; Complete Proteome.
 CC SEQUENCE 54 AA; 6444 MW; 153B7BB801EDD963 CRC64;
 CC -----
 CC Query Match 89.7%; Score 26; DB 1; Length 54;
 CC Best Local Similarity 80.0%; Pred. No. 47;
 CC Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC Qy 1 TKPPR 5
 CC |||||
 CC 4 TKPPK 8
 CC -----
 CC RESULT 18
 CC SECR_MOUSE STANDARD; PRT; 133 AA.
 CC AC Q09535;
 CC DT 01-OCT-1994 (Rel. 30, Created)
 CC DT 01-OCT-1994 (Rel. 30, Last sequence update)
 CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
 CC DE Secretin precursor.
 CC GN SCT.
 CC OS Mus musculus (Mouse).
 CC CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=94234995; PubMed=8179583;
 CC Lan M.S., Kajiyama W., Donadel G., Lu J., Notkins A.L.;
 CC "cDNA sequence and genomic organization of mouse secretin.";
 CC Biochem. Biophys. Res. Commun. 200:1066-1071(1994).
 CC -!- FUNCTION: Stimulates formation of NaHCO(3)-rich pancreatic juice
 CC and secretion of NaHCO(3)-rich bile and inhibits HCl production by
 CC the stomach.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the glucagon family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U07568; AAA18453.1; -;
 CC EMBL; X73580; CAA5982.1; -;
 CC PIR; JC2202; JC2202.
 CC MGD; MGI:99466; Sct.

DR InterPro; IPR000532; Glucagon.
 DR Pfam; PF00123; hormone2; 1.
 DR SMART; SM00070; GLUCA; 1.
 DR PROSITE; PS00260; GLUCAGON; 1.
 KW Glucagon family; Hormone; Amidation;
 Cleavage on pair of basic residues; Signal.
 FT SIGNAL 1 22 BY SIMILARITY.
 FT PEPTIDE 32 58 SECRETIN (BY SIMILARITY).
 FT MOD_RES 58 58 AMIDATION (G-59 PROVIDE AMIDE GROUP).
 SQ SEQUENCE 133 AA; 14914 MW; 9869BCF74CA9709 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 133;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPRP 5
 DB 129 TRPRP 133

RESULT 19

VEGH ORFN2 STANDARD; PRT; 133 AA.

AC P525B4;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Vascular endothelial growth factor homolog precursor.
 GN A2R.

OS Orf virus (strain NZ2) (OV NZ-2).
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Parapoxvirus.
 OX NCBI_TaxID=10259;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94076465; PubMed=8254780;

RA Lytle D.J., Fraser K.M., Fleming S.B., Mercer A.A., Robinson A.J.;
 RT "Homologs of vascular endothelial growth factor are encoded by the
 RT poxvirus orf virus."
 RL Vitol. 68:84-92(1994).

CC -1- FUNCTION: INDUCES ENDOTHELIAL PROLIFERATION.

CC -1- SUBUNIT: Homodimer; disulfide-linked (BY similarity).

CC -1- SIMILARITY: Belongs to the PDGF/VEGF growth factor family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC

DR EMBL; S67520; AAB29220.2; -.

DR HSSP; P15692; 1VPP.

DR InterPro; IPR002400; GF_cysknot.

DR InterPro; IPR000072; PD_growth_factor.

DR Pfam; PF00341; PDGF; 1.

DR PRINTS; PR00438; GFCSKNOT.

DR ProDom; PDC01629; pd_growth_factor; 1.

DR SMART; SM00141; PDGF; 1.

DR PROSITE; PS00249; PDGF_1; 1.

DR PROSITE; PS02078; PDGF_2; 1.

KW Mitogen; Growth factor; Glycoprotein; Signal.

FT SIGNAL 1 20 POTENTIAL.

FT CHAIN 21 133 VASCULAR ENDOTHELIAL GROWTH FACTOR

FT HOMOLOG.

FT DISULFID 36 78 BY SIMILARITY.

FT DISULFID 67 112 BY SIMILARITY.

FT DISULFID 71 114 BY SIMILARITY.

FT DISULFID 61 61 INTERCHAIN (BY SIMILARITY).

FT DISULFID 70 70 INTERCHAIN (BY SIMILARITY).

FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT SEQUENCE 133 AA; 14715 MW; 917C0F6883030C39 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 133;
 Best Local Similarity 80.0%; Pred. No. 1.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPRP 5
 DB 126 TRPRP 130

RESULT 20

RT12 DROME STANDARD; PRT; 140 AA.

AC P10735; Q9V3R2;

DT 01-JUL-1989 (Rel. 11, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE 40S ribosomal protein S12, mitochondrial precursor (MT-RPS12)

DE (technical knockout locus protein).

GN TKO OR EG:BACH5J11.1 OR CG7925.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88027001; PubMed=3117373;

RA Royden C.S., Pirrotta V., Jan L.Y.;

RT "The tko locus, site of a behavioral mutation in D. melanogaster,
 RT codes for a protein homologous to prokaryotic ribosomal protein
 RT S12."

RL Cell 51:165-173(1987).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Oregon-R;

RX MEDLINE=20196011; PubMed=10731137;

RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,

RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demailles J., Cadieu E.,

RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,

RA Papagannakis G., Spanos L., Cox S., Siden-Kiamos I., Bolshakov S.,

RA Modell J., Peter A., Schoettler P., Werner M., Moutkioti F.,

RA Benest N., Dove G., Schaefer U., Jaekle H., Bucheton A.,

RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,

RA McWilliam P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,

RA Glover D.M.;

RT "From sequence to chromosome: the tip of the X chromosome of D.
 RT melanogaster."

RL Science 287:2220-2222(2000).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blasej R.G., Champ M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,

RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Zoller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan Z., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeswam C.,

```
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas P., Testor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
CC Science 287:2185-2195(2000)
CC -!- SUBCELLULAR LOCATION: Mitochondrial.
CC -!- MISCELLANEOUS: MUTATION OF TKO CAUSES BEHAVIORAL MUTATION ("BANG
CC SENSITIVITY" = TEMPORARILY PARALYSIS IN RESPONSE TO A PHYSICAL
CC JOLT).
CC -!- SIMILARITY: Belongs to the S12P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M19494; AAA28935.1; -.
CC EMBL; AL33505; CAB65841.1; -.
CC EMBL; AE003424; AAF45781.1; -.
CC PIR; A29622; A29622.
CC FlyBase; FBgn0003714; tko.
CC GO; GO:0008049; P:male courtship behavior; IMP.
CC GO; GO:0007638; P:mechanosensory behavior; IMP.
CC GO; GO:0009592; P:perception of sound; IMP.
CC GO; GO:00095612; P:response to mechanical stimulus; IMP.
CC InterPro; IPR008994; Nucleic_acid_OB.
CC InterPro; IPR006032; Ribosomal_S12_23.
CC InterPro; IPR005679; Ribosomal_S12B5/c.
CC Pfam; PF00164; Ribosomal_S12; 1.
CC PRINTS; PF01034; RIBOSOMALS12.
CC PRODOM; PD000576; Ribosomal_S12_23; 1.
CC TIGRFAMs; TIGR00981; rpsL_bact; 1.
CC PROSITE; PS00055; RIBOSOMAL_S12; 1.
KW Ribosomal protein; Mitochondrion; Transit peptide.
FT TRANSIT 1 30 MITOCHONDRION (POTENTIAL).
FT CHAIN 31 140 40S RIBOSOMAL PROTEIN S12.
SQ SEQUENCE 140 AA; 15532 MW; 8187A7312C529F31 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 140;
Best Local Similarity 80.0%; Pred.No. 1.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPRP 5
DB 45 TRPRP 49

RESULT 21
KGUA SYNEL
ID KGUA SYNEL STANDARD; PRT; 191 AA.
AC Q8DMQ7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN GMK OR TLL0054.
```

```
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
CX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SP-1;
RX MEDLINE=22255144; PubMed=12240834;
RA Nakamura Y., Kaneo T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus SP-1.";
RL DNA Res. 9:123-130(2002).
CC -!- FUNCTION: Essential for recycling GMP and indirectly, cGMP.
CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the guanylate kinase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP005369; BAC07607.1; -.
CC HAMAP; MF 00328; -.
CC InterPro; IPR008144; Guanylate_kin.
CC InterPro; IPR008145; Guanylt/Ca.
CC Pfam; PF00625; Guanylate_kin; 1.
CC SMART; SM00072; GuKc; 1.
CC PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
CC PROSITE; PS00052; GUANYLATE_KINASE_2; 1.
KW Transferase; Kinase; ATP-binding; Complete proteome.
FT NP BIND 16 23 ATP (BY SIMILARITY).
SQ SEQUENCE 191 AA; 20990 MW; 70B6C15769D73D51 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 191;
Best Local Similarity 80.0%; Pred.No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPRP 5
DB 44 TRPRP 48

RESULT 22
VCO7 ADE04
ID VCO7 ADE04 STANDARD; PRT; 193 AA.
AC Q96831;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Major core protein precursor (Protein VII) (pVII).
GN PVII.
OS Human adenovirus type 4.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
CX NCBI_TaxID=28280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Isolate RI-6;
RA Tarassishin L., Swalowski P.W.S., McIay J., Russell W.C.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; U70921; AAC83411.1; -.
DR InterPro; IPR004912; Adeno_VII.
DR Pfam; PF03228; Adeno_VII; 1.
FT PROPEP 1 24 BY SIMILARITY.
FT CHAIN 25 193 MAJOR CORE PROTEIN.
FT SITE 24 25 CLEAVAGE (BY ADENOVIRUS PROTEASE)
(POTENTIAL).
SQ SEQUENCE 193 AA; 21358 MW; 43137E07DB379DD0 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 193;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 TKPRP 5
|:|:|
Db 188 TRPPR 192

RESULT 23
VC07_ADE02 STANDARD; PRT; 198 AA.
AC P03266; P12542;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Major core protein precursor (Protein VII) (pVII).
GN pVII.
OS Human adenovirus type 2, and
OS Human adenovirus type 5.
OC Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=10515; 28285;
[1]
[2]
[3]
[4]
[5]
[6]
[7]
[8]
[9]
[10]
[11]
[12]
[13]
[14]
[15]
[16]
[17]
[18]
[19]
[20]
[21]
[22]
[23]
[24]
[25]
[26]
[27]
[28]
[29]
[30]
[31]
[32]
[33]
[34]
[35]
[36]
[37]
[38]
[39]
[40]
[41]
[42]
[43]
[44]
[45]
[46]
[47]
[48]
[49]
[50]
[51]
[52]
[53]
[54]
[55]
[56]
[57]
[58]
[59]
[60]
[61]
[62]
[63]
[64]
[65]
[66]
[67]
[68]
[69]
[70]
[71]
[72]
[73]
[74]
[75]
[76]
[77]
[78]
[79]
[80]
[81]
[82]
[83]
[84]
[85]
[86]
[87]
[88]
[89]
[90]
[91]
[92]
[93]
[94]
[95]
[96]
[97]
[98]
[99]
[100]
[101]
[102]
[103]
[104]
[105]
[106]
[107]
[108]
[109]
[110]
[111]
[112]
[113]
[114]
[115]
[116]
[117]
[118]
[119]
[120]
[121]
[122]
[123]
[124]
[125]
[126]
[127]
[128]
[129]
[130]
[131]
[132]
[133]
[134]
[135]
[136]
[137]
[138]
[139]
[140]
[141]
[142]
[143]
[144]
[145]
[146]
[147]
[148]
[149]
[150]
[151]
[152]
[153]
[154]
[155]
[156]
[157]
[158]
[159]
[160]
[161]
[162]
[163]
[164]
[165]
[166]
[167]
[168]
[169]
[170]
[171]
[172]
[173]
[174]
[175]
[176]
[177]
[178]
[179]
[180]
[181]
[182]
[183]
[184]
[185]
[186]
[187]
[188]
[189]
[190]
[191]
[192]
[193]
[194]
[195]
[196]
[197]
[198]
[199]
[200]
[201]
[202]
[203]
[204]
[205]
[206]
[207]
[208]
[209]
[210]
[211]
[212]
[213]
[214]
[215]
[216]
[217]
[218]
[219]
[220]
[221]
[222]
[223]
[224]
[225]
[226]
[227]
[228]
[229]
[230]
[231]
[232]
[233]
[234]
[235]
[236]
[237]
[238]
[239]
[240]
[241]
[242]
[243]
[244]
[245]
[246]
[247]
[248]
[249]
[250]
[251]
[252]
[253]
[254]
[255]
[256]
[257]
[258]
[259]
[260]
[261]
[262]
[263]
[264]
[265]
[266]
[267]
[268]
[269]
[270]
[271]
[272]
[273]
[274]
[275]
[276]
[277]
[278]
[279]
[280]
[281]
[282]
[283]
[284]
[285]
[286]
[287]
[288]
[289]
[290]
[291]
[292]
[293]
[294]
[295]
[296]
[297]
[298]
[299]
[300]
[301]
[302]
[303]
[304]
[305]
[306]
[307]
[308]
[309]
[310]
[311]
[312]
[313]
[314]
[315]
[316]
[317]
[318]
[319]
[320]
[321]
[322]
[323]
[324]
[325]
[326]
[327]
[328]
[329]
[330]
[331]
[332]
[333]
[334]
[335]
[336]
[337]
[338]
[339]
[340]
[341]
[342]
[343]
[344]
[345]
[346]
[347]
[348]
[349]
[350]
[351]
[352]
[353]
[354]
[355]
[356]
[357]
[358]
[359]
[360]
[361]
[362]
[363]
[364]
[365]
[366]
[367]
[368]
[369]
[370]
[371]
[372]
[373]
[374]
[375]
[376]
[377]
[378]
[379]
[380]
[381]
[382]
[383]
[384]
[385]
[386]
[387]
[388]
[389]
[390]
[391]
[392]
[393]
[394]
[395]
[396]
[397]
[398]
[399]
[400]
[401]
[402]
[403]
[404]
[405]
[406]
[407]
[408]
[409]
[410]
[411]
[412]
[413]
[414]
[415]
[416]
[417]
[418]
[419]
[420]
[421]
[422]
[423]
[424]
[425]
[426]
[427]
[428]
[429]
[430]
[431]
[432]
[433]
[434]
[435]
[436]
[437]
[438]
[439]
[440]
[441]
[442]
[443]
[444]
[445]
[446]
[447]
[448]
[449]
[450]
[451]
[452]
[453]
[454]
[455]
[456]
[457]
[458]
[459]
[460]
[461]
[462]
[463]
[464]
[465]
[466]
[467]
[468]
[469]
[470]
[471]
[472]
[473]
[474]
[475]
[476]
[477]
[478]
[479]
[480]
[481]
[482]
[483]
[484]
[485]
[486]
[487]
[488]
[489]
[490]
[491]
[492]
[493]
[494]
[495]
[496]
[497]
[498]
[499]
[500]
[501]
[502]
[503]
[504]
[505]
[506]
[507]
[508]
[509]
[510]
[511]
[512]
[513]
[514]
[515]
[516]
[517]
[518]
[519]
[520]
[521]
[522]
[523]
[524]
[525]
[526]
[527]
[528]
[529]
[530]
[531]
[532]
[533]
[534]
[535]
[536]
[537]
[538]
[539]
[540]
[541]
[542]
[543]
[544]
[545]
[546]
[547]
[548]
[549]
[550]
[551]
[552]
[553]
[554]
[555]
[556]
[557]
[558]
[559]
[560]
[561]
[562]
[563]
[564]
[565]
[566]
[567]
[568]
[569]
[570]
[571]
[572]
[573]
[574]
[575]
[576]
[577]
[578]
[579]
[580]
[581]
[582]
[583]
[584]
[585]
[586]
[587]
[588]
[589]
[590]
[591]
[592]
[593]
[594]
[595]
[596]
[597]
[598]
[599]
[600]
[601]
[602]
[603]
[604]
[605]
[606]
[607]
[608]
[609]
[610]
[611]
[612]
[613]
[614]
[615]
[616]
[617]
[618]
[619]
[620]
[621]
[622]
[623]
[624]
[625]
[626]
[627]
[628]
[629]
[630]
[631]
[632]
[633]
[634]
[635]
[636]
[637]
[638]
[639]
[640]
[641]
[642]
[643]
[644]
[645]
[646]
[647]
[648]
[649]
[650]
[651]
[652]
[653]
[654]
[655]
[656]
[657]
[658]
[659]
[660]
[661]
[662]
[663]
[664]
[665]
[666]
[667]
[668]
[669]
[670]
[671]
[672]
[673]
[674]
[675]
[676]
[677]
[678]
[679]
[680]
[681]
[682]
[683]
[684]
[685]
[686]
[687]
[688]
[689]
[690]
[691]
[692]
[693]
[694]
[695]
[696]
[697]
[698]
[699]
[700]
[701]
[702]
[703]
[704]
[705]
[706]
[707]
[708]
[709]
[710]
[711]
[712]
[713]
[714]
[715]
[716]
[717]
[718]
[719]
[720]
[721]
[722]
[723]
[724]
[725]
[726]
[727]
[728]
[729]
[730]
[731]
[732]
[733]
[734]
[735]
[736]
[737]
[738]
[739]
[740]
[741]
[742]
[743]
[744]
[745]
[746]
[747]
[748]
[749]
[750]
[751]
[752]
[753]
[754]
[755]
[756]
[757]
[758]
[759]
[760]
[761]
[762]
[763]
[764]
[765]
[766]
[767]
[768]
[769]
[770]
[771]
[772]
[773]
[774]
[775]
[776]
[777]
[778]
[779]
[780]
[781]
[782]
[783]
[784]
[785]
[786]
[787]
[788]
[789]
[790]
[791]
[792]
[793]
[794]
[795]
[796]
[797]
[798]
[799]
[800]
[801]
[802]
[803]
[804]
[805]
[806]
[807]
[808]
[809]
[810]
[811]
[812]
[813]
[814]
[815]
[816]
[817]
[818]
[819]
[820]
[821]
[822]
[823]
[824]
[825]
[826]
[827]
[828]
[829]
[830]
[831]
[832]
[833]
[834]
[835]
[836]
[837]
[838]
[839]
[840]
[841]
[842]
[843]
[844]
[845]
[846]
[847]
[848]
[849]
[850]
[851]
[852]
[853]
[854]
[855]
[856]
[857]
[858]
[859]
[860]
[861]
[862]
[863]
[864]
[865]
[866]
[867]
[868]
[869]
[870]
[871]
[872]
[873]
[874]
[875]
[876]
[877]
[878]
[879]
[880]
[881]
[882]
[883]
[884]
[885]
[886]
[887]
[888]
[889]
[890]
[891]
[892]
[893]
[894]
[895]
[896]
[897]
[898]
[899]
[900]
[901]
[902]
[903]
[904]
[905]
[906]
[907]
[908]
[909]
[910]
[911]
[912]
[913]
[914]
[915]
[916]
[917]
[918]
[919]
[920]
[921]
[922]
[923]
[924]
[925]
[926]
[927]
[928]
[929]
[930]
[931]
[932]
[933]
[934]
[935]
[936]
[937]
[938]
[939]
[940]
[941]
[942]
[943]
[944]
[945]
[946]
[947]
[948]
[949]
[950]
[951]
[952]
[953]
[954]
[955]
[956]
[957]
[958]
[959]
[960]
[961]
[962]
[963]
[964]
[965]
[966]
[967]
[968]
[969]
[970]
[971]
[972]
[973]
[974]
[975]
[976]
[977]
[978]
[979]
[980]
[981]
[982]
[983]
[984]
[985]
[986]
[987]
[988]
[989]
[990]
[991]
[992]
[993]
[994]
[995]
[996]
[997]
[998]
[999]
[1000]
```

```
DR EMBL; M73260; AAA96408.1; ALT_SEQ.
DR EMBL; M22141; AAA42520.1; -.
DR PIR; C03837; FOAD72.
DR PIR; P0067; FOAD72.
DR InterPro; IPR004912; Adeno_VII.
DR Pfam; PF03228; Adeno_VII; 1.
KW Core protein; Late protein.
FT PROPEP 1 24
FT CHAIN 25 198 MAJOR CORE PROTEIN.
FT SITE 24 25 CLEAVAGE (BY ADENOVIRUS PROTEASE).
FT CONFLICT 112 112 MISSING (IN REF. 2).
SQ SEQUENCE 198 AA; 21992 MW; 7D5A8D426F08E952 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 198;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Gaps 0;

Qy 1 TKPRP 5
|:|:|
Db 193 TRPPR 197

RESULT 24
KGUA_ANASP STANDARD; PRT; 199 AA.
ID KGUA_ANASP
AC Q8Z017;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN GMP_KR_ALR0106.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostoc.
OX NCBI_TaxID=103690;
[1]
[2]
[3]
[4]
[5]
[6]
[7]
[8]
[9]
[10]
[11]
[12]
[13]
[14]
[15]
[16]
[17]
[18]
[19]
[20]
[21]
[22]
[23]
[24]
[25]
[26]
[27]
[28]
[29]
[30]
[31]
[32]
[33]
[34]
[35]
[36]
[37]
[38]
[39]
[40]
[41]
[42]
[43]
[44]
[45]
[46]
[47]
[48]
[49]
[50]
[51]
[52]
[53]
[54]
[55]
[56]
[57]
[58]
[59]
[60]
[61]
[62]
[63]
[64]
[65]
[66]
[67]
[68]
[69]
[70]
[71]
[72]
[73]
[74]
[75]
[76]
[77]
[78]
[79]
[80]
[81]
[82]
[83]
[84]
[85]
[86]
[87]
[88]
[89]
[90]
[91]
[92]
[93]
[94]
[95]
[96]
[97]
[98]
[99]
[100]
[101]
[102]
[103]
[104]
[105]
[106]
[107]
[108]
[109]
[110]
[111]
[112]
[113]
[114]
[115]
[116]
[117]
[118]
[119]
[120]
[121]
[122]
[123]
[124]
[125]
[126]
[127]
[128]
[129]
[130]
[131]
[132]
[133]
[134]
[135]
[136]
[137]
[138]
[139]
[140]
[141]
[142]
[143]
[144]
[145]
[146]
[147]
[148]
[149]
[150]
[151]
[152]
[153]
[154]
[155]
[156]
[157]
[158]
[159]
[160]
[161]
[162]
[163]
[164]
[165]
[166]
[167]
[168]
[169]
[170]
[171]
[172]
[173]
[174]
[175]
[176]
[177]
[178]
[179]
[180]
[181]
[182]
[183]
[184]
[185]
[186]
[187]
[188]
[189]
[190]
[191]
[192]
[193]
[194]
[195]
[196]
[197]
[198]
[199]
[200]
[201]
[202]
[203]
[204]
[205]
[206]
[207]
[208]
[209]
[210]
[211]
[212]
[213]
[214]
[215]
[216]
[217]
[218]
[219]
[220]
[221]
[222]
[223]
[224]
[225]
[226]
[227]
[228]
[229]
[230]
[231]
[232]
[233]
[234]
[235]
[236]
[237]
[238]
[239]
[240]
[241]
[242]
[243]
[244]
[245]
[246]
[247]
[248]
[249]
[250]
[251]
[252]
[253]
[254]
[255]
[256]
[257]
[258]
[259]
[260]
[261]
[262]
[263]
[264]
[265]
[266]
[267]
[268]
[269]
[270]
[271]
[272]
[273]
[274]
[275]
[276]
[277]
[278]
[279]
[280]
[281]
[282]
[283]
[284]
[285]
[286]
[287]
[288]
[289]
[290]
[291]
[292]
[293]
[294]
[295]
[296]
[297]
[298]
[299]
[300]
[301]
[302]
[303]
[304]
[305]
[306]
[307]
[308]
[309]
[310]
[311]
[312]
[313]
[314]
[315]
[316]
[317]
[318]
[319]
[320]
[321]
[322]
[323]
[324]
[325]
[326]
[327]
[328]
[329]
[330]
[331]
[332]
[333]
[334]
[335]
[336]
[337]
[338]
[339]
[340]
[341]
[342]
[343]
[344]
[345]
[346]
[347]
[348]
[349]
[350]
[351]
[352]
[353]
[354]
[355]
[356]
[357]
[358]
[359]
[360]
[361]
[362]
[363]
[364]
[365]
[366]
[367]
[368]
[369]
[370]
[371]
[372]
[373]
[374]
[375]
[376]
[377]
[378]
[379]
[380]
[381]
[382]
[383]
[384]
[385]
[386]
[387]
[388]
[389]
[390]
[391]
[392]
[393]
[394]
[395]
[396]
[397]
[398]
[399]
[400]
[401]
[402]
[403]
[404]
[405]
[406]
[407]
[408]
[409]
[410]
[411]
[412]
[413]
[414]
[415]
[416]
[417]
[418]
[419]
[420]
[421]
[422]
[423]
[424]
[425]
[426]
[427]
[428]
[429]
[430]
[431]
[432]
[433]
[434]
[435]
[436]
[437]
[438]
[439]
[440]
[441]
[442]
[443]
[444]
[445]
[446]
[447]
[448]
[449]
[450]
[451]
[452]
[453]
[454]
[455]
[456]
[457]
[458]
[459]
[460]
[461]
[462]
[463]
[464]
[465]
[466]
[467]
[468]
[469]
[470]
[471]
[472]
[473]
[474]
[475]
[476]
[477]
[478]
[479]
[480]
[481]
[482]
[483]
[484]
[485]
[486]
[487]
[488]
[489]
[490]
[491]
[492]
[493]
[494]
[495]
[496]
[497]
[498]
[499]
[500]
[501]
[502]
[503]
[504]
[505]
[506]
[507]
[508]
[509]
[510]
[511]
[512]
[513]
[514]
[515]
[516]
[517]
[518]
[519]
[520]
[521]
[522]
[523]
[524]
[525]
[526]
[527]
[528]
[529]
[530]
[531]
[532]
[533]
[534]
[535]
[536]
[537]
[538]
[539]
[540]
[541]
[542]
[543]
[544]
[545]
[546]
[547]
[548]
[549]
[550]
[551]
[552]
[553]
[554]
[555]
[556]
[557]
[558]
[559]
[560]
[561]
[562]
[563]
[564]
[565]
[566]
[567]
[568]
[569]
[570]
[571]
[572]
[573]
[574]
[575]
[576]
[577]
[578]
[579]
[580]
[581]
[582]
[583]
[584]
[585]
[586]
[587]
[588]
[589]
[590]
[591]
[592]
[593]
[594]
[595]
[596]
[597]
[598]
[599]
[600]
[601]
[602]
[603]
[604]
[605]
[606]
[607]
[608]
[609]
[610]
[611]
[612]
[613]
[614]
[615]
[616]
[617]
[618]
[619]
[620]
[621]
[622]
[623]
[624]
[625]
[626]
[627]
[628]
[629]
[630]
[631]
[632]
[633]
[634]
[635]
[636]
[637]
[638]
[639]
[640]
[641]
[642]
[643]
[644]
[645]
[646]
[647]
[648]
[649]
[650]
[651]
[652]
[653]
[654]
[655]
[656]
[657]
[658]
[659]
[660]
[661]
[662]
[663]
[664]
[665]
[666]
[667]
[668]
[669]
[670]
[671]
[672]
[673]
[674]
[675]
[676]
[677]
[678]
[679]
[680]
[681]
[682]
[683]
[684]
[685]
[686]
[687]
[688]
[689]
[690]
[691]
[692]
[693]
[694]
[695]
[696]
[697]
[698]
[699]
[700]
[701]
[702]
[703]
[704]
[705]
[706]
[707]
[708]
[709]
[710]
[711]
[712]
[713]
[714]
[715]
[716]
[717]
[718]
[719]
[720]
[721]
[722]
[723]
[724]
[725]
[726]
[727]
[728]
[729]
[730]
[731]
[732]
[733]
[734]
[735]
[736]
[737]
[738]
[739]
[740]
[741]
[742]
[743]
[744]
[745]
[746]
[747]
[748]
[749]
[750]
[751]
[752]
[753]
[754]
[755]
[756]
[757]
[758]
[759]
[760]
[761]
[762]
[763]
[764]
[765]
[766]
[767]
[768]
[769]
[770]
[771]
[772]
[773]
[774]
[775]
[776]
[777]
[778]
[779]
[780]
[781]
[782]
[783]
[784]
[785]
[786]
[787]
[788]
[789]
[790]
[791]
[792]
[793]
[794]
[795]
[796]
[797]
[798]
[799]
[800]
[801]
[802]
[803]
[804]
[805]
[806]
[807]
[808]
[809]
[810]
[811]
[812]
[813]
[814]
[815]
[816]
[817]
[818]
[819]
[820]
[821]
[822]
[823]
[824]
[825]
[826]
[827]
[828]
[829]
[830]
[831]
[832]
[833]
[834]
[835]
[836]
[837]
[838]
[839]
[840]
[841]
[842]
[843]
[844]
[845]
[846]
[847]
[848]
[849]
[850]
[851]
[852]
[853]
[854]
[855]
[856]
[857]
[858]
[859]
[860]
[861]
[862]
[863]
[864]
[865]
[866]
[867]
[868]
[869]
[870]
[871]
[872]
[873]
[874]
[875]
[876]
[877]
[878]
[879]
[880]
[881]
[882]
[883]
[884]
[885]
[886]
[887]
[888]
[889]
[890]
[891]
[892]
[893]
[894]
[895]
[896]
[897]
[898]
[899]
[900]
[901]
[902]
[903]
[904]
[905]
[906]
[907]

```

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
|:|:|

Db 55 TRPPR 59

RESULT 25

KGUA_NEIMA
ID KGUA_NEIMA STANDARD; PRT; 205 AA.

AC QJ756;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN GMP OR NMA1919.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / Serogroup A / Serotype 4A;
RC MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Moralli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Felwell T., Hamlin N., Holroyd S.,
RA Jørgens K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491";
RL Nature 404:502-506 (2000).
CC -!- FUNCTION: Essential for recycling GMP and indirectly, cGMP.
CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the guanylate kinase family.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AL162757; CAB85140.1; -
FIR; G81819; G81819.
HSP; P15454; IGKY.
HMAP; MF_00328; -; 1.
InterPro; IPR008144; Guanylate_kin.
InterPro; IPR008145; Guanylt/Ca.
Pfam; PF00625; Guanylate_kin; 1.
SMART; SM00072; GuKc; 1.
PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
Transferase; Kinase; ATP-binding; Complete proteome.
NP_BIND 14 21 ATP (BY SIMILARITY).
SQ SEQUENCE 205 AA; 22530 MW; 04FA07E5450C9007 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 205;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

|:|:|

Db 42 TRPPR 46

RESULT 26

KGUA_NEIMB
ID KGUA_NEIMB STANDARD; PRT; 205 AA.
AC Q9JYB5;

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Guanylate kinase (EC 2.7.4.8) (GMP kinase).
GN GMP OR NMB1661.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RC MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Ciftone H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizzi M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58";
RL Science 287:1809-1815 (2000).
CC -!- FUNCTION: Essential for recycling GMP and indirectly, cGMP.
CC -!- CATALYTIC ACTIVITY: ATP + GMP = ADP + GDP.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the guanylate kinase family.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to license@isb-sib.ch).

EMBL; AE002517; AAF42010.1; -
FIR; G81055; G81055.
HSP; P15454; IGKY.
HMAP; NMB1661; -; 1.
HMAP; MF_00328; -; 1.
InterPro; IPR008144; Guanylate_kin.
InterPro; IPR008145; Guanylt/Ca.
Pfam; PF00625; Guanylate_kin; 1.
SMART; SM00072; GuKc; 1.
PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
PROSITE; PS50052; GUANYLATE_KINASE_2; 1.
Transferase; Kinase; ATP-binding; Complete proteome.
NP_BIND 14 21 ATP (BY SIMILARITY).
SQ SEQUENCE 205 AA; 22500 MW; 54EE07E545189008 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 205;
Best Local Similarity 80.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
|:|:|

Db 42 TRPPR 46

RESULT 27
Y132_NPVAC STANDARD; PRT; 219 AA.
ID Y132_NPVAC
AC P24730;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Hypothetical 25.1 kDa protein in pP34-EXO intergenic region (ORF 4).
OS Autographa californica nuclear polyhedrosis virus (AcMNPV).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae;
OC Nucleopolyhedrovirus.
OX NCBI_TaxID=46015;

```

RV RP SEQUENCE FROM N.A.
RX MEDLINE=97311863; PubMed=3041026;
RA Oellig C., Happ B., Mueller T., Doerfler W.;
RT "Overlapping sets of viral RNAs reflect the array of polypeptides in
RT the SCOBI J and N fragments (map positions 81.2 to 85.0) of the
RT Autographa californica nuclear polyhedrosis virus genome.";
RL J. Virol. 61:3048-3057(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C6;
RX MEDLINE=94303173; PubMed=8030224;
RA Ayres M.D., Howard S.C., Kuzio J., Lopez-Ferber M., Possee R.D.;
RT "The complete DNA sequence of Autographa californica nuclear
RT polyhedrosis virus."
RL Virology 202:586-605(1994).
CC -1- SIMILARITY: TO CORRESPONDING ORF IN OPNVPV.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; ML7548; AAA66806.1; -.
CC EMBL; L22858; AAA66762.1; -.
CC PIR; E72866; E72866.
CC KW Hypothetical protein; Late protein.
CC SEQUENCE 219 AA; 25136 MW; 281E1625BE8F6A5E CRC64;
CC
CC
CC Query Match 89.7%; Score 26; DB 1; Length 219;
CC Best Local Similarity 80.0%; Pred. No. 2e+02;
CC Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 TKPPR 5
CC Db 23 TKPPK 27
CC
CC RESULT 28
CC PESS_LUCCU
CC ID PESS_LUCCU STANDARD; PRT; 220 AA.
CC AC Q95UE8; Q8MUP5;
CC DT 10-OCT-2003 (Rel. 42, Created)
CC DT 10-OCT-2003 (Rel. 42, Last sequence update)
CC DT 15-MAR-2004 (Rel. 43, Last annotation update)
CC DE Peritrophin-55 precursor.
CC OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).
CC OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
CC OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
CC OC Calliphoridae; Lucilia.
CC OX NCBI_TaxID=7375;
CC RN [1]
CC RP SEQUENCE FROM N.A.; SEQUENCE OF 20-57 AND 60-94, TISSUE SPECIFICITY,
CC AND DEVELOPMENTAL STAGE.
CC RC TISSUE=Larva;
CC RA MEDLINE=22423280; PubMed=12535682;
CC RA Tellam R.L., Vuocolo T., Eisemann C.H., Briscoe S., Riding G.A.,
CC RA Elvin C.N., Pearson R.D.;
CC RT "Identification of an immuno-protective mucin-like protein,
CC RT peritrophin-55, from the peritrophic matrix of Lucilia cuprina
CC RT larvae."
CC RL Insect Biochem. Mol. Biol. 33:239-252(2003).
CC CC -1- FUNCTION: May bind oligosaccharide structures.
CC CC -1- TISSUE SPECIFICITY: Larval peritrophic membrane.
CC CC -1- DEVELOPMENTAL STAGE: Expressed in all 3 larval instars but not
CC CC adults or eggs.
CC CC -1- PTM: Glycosylated.
CC CC -1- SIMILARITY: Contains 1 chitin-binding type-2 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY055470; AAL15463.1; -.
CC EMBL; AF515826; AAM55223.1; -.
CC InterPro; IPR002557; Chitin bind PerA.
CC PROSITE; PS0940; CHIT_BIND_II; I.
CC KW Glycoprotein; Signal.
CC FT SIGNAL 1 19
CC FT CHAIN 20 220
CC FT DOMAIN 33 95
CC FT CARBOHYD 29 29
CC FT CONFLICT 67 67
CC FT CONFLICT 69 69
CC FT CONFLICT 142 142
CC FT CONFLICT 142 142
CC FT CONFLICT 164 164
CC FT CONFLICT 190 190
CC FT CONFLICT 190 190
CC SEQUENCE 220 AA; 23535 MW; 88C74ED57F2ED7C7 CRC64;
CC
CC Query Match 89.7%; Score 26; DB 1; Length 220;
CC Best Local Similarity 80.0%; Pred. No. 2e+02;
CC Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 TKPPR 5
CC Db 155 TKPPK 159
CC
CC RESULT 29
CC NOCT_RAT
CC ID NOCT_RAT STANDARD; PRT; 253 AA.
CC AC Q9ET55;
CC DT 28-FEB-2003 (Rel. 41, Created)
CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Nocturnin (CCR4 protein homolog) (Fragment).
CC GN CCRN4L OR NOC OR CCR4.
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC OX NCBI_TaxID=10116;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE=Retina;
CC RA Wang Y., Osterbur D.L., Green C.B., Besharse J.C.;
CC RT "Mammalian homologs of Xenopus nocturnin: conservation of structure
CC RT and circadian regulation."
CC RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC CC -1- FUNCTION: Component of the circadian clock or downstream effector
CC CC of clock function. Exhibits a high amplitude circadian rhythm with
CC CC maximal levels in early evening. In constant darkness or constant
CC CC light, the amplitude of the rhythm decreases (By similarity).
CC CC -1- SIMILARITY: Belongs to the CCR4/nocturnin family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF199495; ARG01390.1; -.
CC InterPro; IPR005135; Exo endo_phos.
CC Pfam; PF03372; Exo_endo_phos; 1.
CC KW Biological rhythms.
CC FT NON_TER 1 1
CC FT NON_TER 253 253
CC SEQUENCE 253 AA; 28662 MW; C63565735BCC432E CRC64;

```

```
Query Match      89.7%; Score 26; DB 1; Length 253;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPRP 5
   |||||
Db 44 TRPRP 48

RESULT 30
TRUA_XYLFA STANDARD; PRT; 257 AA.
AC Q9PDK6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE tRNA pseudouridine synthase A (EC 4.2.1.70) (Pseudouridylylate synthase
DE I) (Pseudouridine synthase I) (Uracil hydrolyase).
GN TRUA OR XF1373.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=945C;
RX Simpson A.J.G., Reinach P.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvares R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.B.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hdhisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.B., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
CC -!- FUNCTION: Formation of pseudouridine at positions 38, 39 and 40 in
CC the anticodon stem and loop of transfer RNAs (By similarity).
CC -!- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine
CC 5'-phosphate + H(2)O.
CC -!- SIMILARITY: Belongs to the pseudouridine synthase trUA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE012555; AAC28482.1; ALT_INIT.
CC HAMAP; MF_00171; -.
CC InterPro; IPR001406; PseudoU synth_1.
CC Pfam; PF01416; Pseudou synth_1; 2.
CC TIGRFAMs; TIGR00071; hsf_truA; 1.
CC Lyase; tRNA processing; Complete proteome.
CC ACT SITE 53 53 BY SIMILARITY.
CC SEQUENCE 257 AA; 28664 MW; F6D5A76D9C2A88EC CRC64;

DR InterPro; IPR001406; Pseudou synth_1.
DR Pfam; PF01416; Pseudou synth_1; 2.
DR TIGRFAMs; TIGR00071; hsf_truA; 1.
DR Lyase; tRNA processing; Complete proteome.
FT ACT_SITE 53 53 BY SIMILARITY.
SQ SEQUENCE 257 AA; 28638 MW; C717E0D287C7D3F3 CRC64;

Query Match      89.7%; Score 26; DB 1; Length 257;
Best Local Similarity 80.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPRP 5
   |||||
Db 70 TRPRP 74

RESULT 31
TRUA_XYLFT STANDARD; PRT; 257 AA.
AC Q87DS1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE tRNA pseudouridine synthase A (EC 4.2.1.70) (Pseudouridylylate synthase
DE I) (Pseudouridine synthase I) (Uracil hydrolyase).
GN TRUA OR PD0610.
OS Xylella fastidiosa (strain Temecula / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=22421331; PubMed=12533478;
RX Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B., Moon D.H.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., da Silva F.R.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.V.F., El-Dorry H., Tsai S.M.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., Nunes L.R., Siqueira W.J.,
RA Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Kuramae E.B.,
RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Amaral A.M.,
RA Marino C.S., Gaglioti E., Abreu L.L., Alves L.M.C., Celestino A.V.,
RA Baia G.S., Blanco S.R., Brito W.S., Camargo F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Tsukumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.P.;
RX "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa.";
RL J. Bacteriol. 185:1018-1026(2003).
CC -!- FUNCTION: Formation of pseudouridine at positions 38, 39 and 40 in
CC the anticodon stem and loop of transfer RNAs (By similarity).
CC -!- CATALYTIC ACTIVITY: Uracil + D-ribose 5-phosphate = pseudouridine
CC 5'-phosphate + H(2)O.
CC -!- SIMILARITY: Belongs to the pseudouridine synthase trUA family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE012555; AAC28482.1; ALT_INIT.
CC HAMAP; MF_00171; -.
CC InterPro; IPR001406; PseudoU synth_1.
CC Pfam; PF01416; Pseudou synth_1; 2.
CC TIGRFAMs; TIGR00071; hsf_truA; 1.
CC Lyase; tRNA processing; Complete proteome.
CC ACT SITE 53 53 BY SIMILARITY.
CC SEQUENCE 257 AA; 28664 MW; F6D5A76D9C2A88EC CRC64;
```

Query Match 89.7%; Score 26; DB 1; Length 257;
 Best Local Similarity 80.0%; Pred. No. 2.3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 70 TRPPR 74

RESULT 32
 YCOB_CABEL STANDARD; PRT; 304 AA.
 AC Q09360;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Hypothetical 33.0 kDa protein EED8.11 in chromosome II precursor.
 GN EED8.11.
 OS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 CC Rhabditidae; Pelodierinae; Caenorhabditis.
 CC NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Chissee S.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: SOME, TO C.ELEGANS R13F6.2 AND R13F6.8.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U23484; AAC46771.1; -;
 DR PIR; T15922; T15922.
 DR WormPep; EED8.11; C01884.
 DR InterPro; IPR001304; Lectin_C.
 DR SMART; SM00034; CLEC1.1.
 KW Hypothetical protein; Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 304 HYPOTHETICAL PROTEIN EED8.11.
 FT DOMAIN 64 92 POLY-THR.
 SQ SEQUENCE 304 AA; 32982 MW; 60C223B88F534151 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 304;
 Best Local Similarity 80.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 51 TKPPK 55

RESULT 33
 YB04_AQUAE STANDARD; PRT; 350 AA.
 AC Q67189;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein AQ_1104.
 GN AQ_1104.
 OS Aquifex aeolicus.
 CC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 CC NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aufay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus.";
 RL Nature 392:353-358(1998).
 CC -1- SIMILARITY: SOME, TO R.PROWAZEKII RP189.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AE000723; AAC07153.1; -;
 DR PIR; B70395; B70395.
 DR InterPro; IPR008921; Pol_clamp_load_C.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 350 AA; 40693 MW; 2949E786E1DAC2F9 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 350;
 Best Local Similarity 80.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 14 TKPPK 18

RESULT 34
 TRMU_RICCN STANDARD; PRT; 370 AA.
 AC Q92110;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
 DE (EC 2.1.1.61).
 GN TRMU OR RC0410.
 OS Rickettsia conorii.
 CC Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
 CC Rickettsiaceae; Rickettsia.
 CC NCBI_TaxID=781;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Malish 7;
 RX MEDLINE=21442074; PubMed=11557893;
 RA Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
 RA Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
 RA Raoult D.;
 RT "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
 RL Science 293:2093-2098(2001).
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
 CC homocysteine + tRNA containing 5-methylaminomethyl-2-
 CC thiouridylate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: Belongs to the trmU family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AE008605; AAL02948.1; -;
 DR PIR; B97751; B97751.
 DR HAMAP; MF_00144; -; 1.
 DR InterPro; IPR004506; TrmU.
 DR Pfam; PF03054; tRNA_Me_trans; 1.


```
DR TIGRFAMS; TIGR00420; trmU; 1.
KW Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ SEQUENCE 370 AA; 40915 MW; 1884AC815E730CE3 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 370;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 318 TKPPR 322

RESULT 35
RT31 HUMAN
ID RT31 HUMAN STANDARD; PRT; 395 AA.
AC Q92655; Q8WTV8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 28S ribosomal protein S31, mitochondrial precursor (S31mt) (MRP-S31)
DE (Imogen 38).
GN MRPS31 OR IMOGN38.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Hutton J.C., Reep B.O.;
RT "Human Imogen 38. T-cell and antibody responses in newly diagnosed
RT diabetic subjects.";
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.W., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Matsushita K., Farmer A.A., Rubin G.M., Hong L.,
RA Brownstein M.J., Soares M.B., Bonaldi M.P., Casavant T.L., Scheetz T.E.,
RA Rana S.S., Loquellano N.A., Peters G.J., Carvanci P., Prange C.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Touchman J.W., Green B.D., Dickson M.C.,
RA Blakesley R.W., Grimwood J., Schmutz J., Myers R.M.,
RA Rodriguez A.C., Blackburn K., Blackburn K., Moseley A., Sprenulli L.L.;
RA Koc E.C., Burkhardt W., Blackburn K., Moseley A., Sprenulli L.L.;
RT "The small subunit of the mammalian mitochondrial ribosome:
RT identification of the full complement of ribosomal proteins present.";
RL J. Biol. Chem. 276:19363-19374(2001).
CC -1- SUBUNIT: Component of the mitochondrial ribosome small subunit
CC (28S) which comprises a 12S rRNA and about 30 distinct proteins.
CC -1- SUBCELLULAR LOCATION: Mitochondrial.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF003000; BAB49856.1; ALT_INIT.
CC HMAP; MF_00144; -; 1.
CC InterPro; IPR004506; TrmU.
CC Pfam; PF03054; tRNA_We_trans; 1.
CC TIGRFAMS; TIGR00420; trmU; 1.
KW Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ SEQUENCE 396 AA; 42537 MW; 66F118AA97E3086 CRC64;

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; 268747; CAA92951.1; -.
CC EMBL; BC022045; AAH22045.1; -.
CC Genew; HGNC:16632; MRPS31.
CC GO; GO:0005739; C:mitochondrion; TAS.
CC Ribosomal protein; Mitochondrion; Transit peptide.
FT TRANSIT 1 65 MITOCHONDRION (POTENTIAL).
FT CHAIN 66 395 28S RIBOSOMAL PROTEIN S31.
FT CONFLICT 80 80 I -> T (IN REF. 1).
FT CONFLICT 132 132 A -> T (IN REF. 1).
FT CONFLICT 279 279 D -> N (IN REF. 1).
SQ SEQUENCE 395 AA; 45300 MW; E9410F46C94C6F3D CRC64;

Query Match 89.7%; Score 26; DB 1; Length 395;
Best Local Similarity 80.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 118 TKPPK 122

RESULT 36
TRMU RHIL0
ID TRMU RHIL0 STANDARD; PRT; 396 AA.
AC Q98H10;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable tRNA (5-methylaminomethyl-2-thiouridyate)-methyltransferase
DE (EC 2.1.1.61).
GN TRMU OR MLR2824.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=MAF303099; PubMed=11214968;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Yochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
CC homocysteine + tRNA containing 5-methylaminomethyl-2-
CC thiouridyate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the trmU family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF003000; BAB49856.1; ALT_INIT.
CC HMAP; MF_00144; -; 1.
CC InterPro; IPR004506; TrmU.
CC Pfam; PF03054; tRNA_We_trans; 1.
CC TIGRFAMS; TIGR00420; trmU; 1.
KW Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ SEQUENCE 396 AA; 42537 MW; 66F118AA97E3086 CRC64;
```

Query Match 89.7%; Score 26; DB 1; Length 396;
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 320 TRPPR 324

RESULT 37
 TRMU_BRUVE STANDARD; PRT; 398 AA.
 AC Q8Y1L6;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
 DE (EC 2.1.1.61).
 GN TRMU OR BMEI0428.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=1756886;
 RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Mijer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Renik G.,
 RA Jablonka L., Larsen M., P'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
 RA Haselkorn R., Kyprides N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
 CC homocysteine + tRNA containing 5-methylaminomethyl-2-
 CC thiouridylate.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the trmu family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AF3305; AF3305.
 DR HAMAP; MF_001444; -.
 DR InterPro; IPR004506; Trmu.
 DR Pfam; PF03054; tRNA.Me.trans; 1.
 DR TIGRFAMs; TIGR00420; trmu; 1.
 KW Transferase; Methyltransferase; tRNA processing; Complete proteome.
 SQ SEQUENCE 398 AA; 43175 MW; 93687D1D8B9ED3 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 398;
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 322 TRPPR 326

RESULT 38
 TRMU_BRUSU STANDARD; PRT; 398 AA.
 AC Q8CY38;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
 DE (EC 2.1.1.61).
 GN TRMU OR BR1591.
 OS Brucella suis.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Brucellaceae; Brucella.
 OX NCBI_TaxID=29461;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1330 / Biovar 1;
 RX MEDLINE=22247741; PubMed=12271122;
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
 RA Read T.D., Dodson R.J., Unayam L., Brinkac L.M., Beanan M.J.,
 RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
 RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E.,
 RA Riedmuller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
 RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.;
 RT "The Brucella suis genome reveals fundamental similarities between
 RT animal and plant pathogens and symbionts";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
 CC homocysteine + tRNA containing 5-methylaminomethyl-2-
 CC thiouridylate.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Belongs to the trmu family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL; AE014453; AAN30496.1; -.
 DR TIGR; BR1591; -.
 DR HAMAP; MF_001444; -.
 DR InterPro; IPR004506; Trmu.
 DR Pfam; PF03054; tRNA.Me.trans; 1.
 DR TIGRFAMs; TIGR00420; trmu; 1.
 KW Transferase; Methyltransferase; tRNA processing; Complete proteome.
 SQ SEQUENCE 398 AA; 43228 MW; 937367AC3B9ED3 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 398;
 Best Local Similarity 80.0%; Pred. No. 3.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 322 TRPPR 326

RESULT 39
 VGLD_HSVEA STANDARD; PRT; 402 AA.
 ID VGLD_HSVEA
 AC P24872;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DE 15-OCT-2001 (Rel. 40, Last annotation update)
 DE Glycoprotein D precursor (Glycoprotein 17/18).
 GN GD OR GP17/18 OR 72.
 OS Equine herpesvirus type 1 (strain AB1) (EHV-1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10328;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=92268882; PubMed=1316942;
 RX Eiton D.M., Halliburton I.W., Killington R.A., Meredith D.M.,
 RA Bonas W.A.;
 RT "Identification of the equine herpesvirus type 1 glycoprotein 17/18
 RT as a homologue of herpes simplex virus glycoprotein D";

J. Gen. Virol. 73:1227-1233(1992).

[2]
RN SEQUENCE OF 242-402 FROM N.A.
RP MEDLINE=91276272; PubMed=1647359;
RX Elton D.M., Halliburton I.W., Killington R.A., Meredith D.M.,
RA Bonas W.A.,
RT "Sequence analysis of the 4.7-Kb BamHI-EcoRI fragment of the equine
RT herpesvirus type-1 short unique region.",
RL Gene 101:203-208(1991).

[4]
RN -1- SIMILARITY: Belongs to the herpesviruses glycoprotein D family.
RP
RC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

[5]
RN EMBL; M60946; AAA46087.1; -
RP EMBL; M36299; AAA66546.1; -
RC InterPro; IPR002896; Herpes glycop D.
RT InterPro; IPR007110; Ig-like.
RF Pfam; PF01537; Herpes glycop D; 1.
RW Glycoprotein; Transmembrane; Signal.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 402 GLYCOPROTEIN D.
FT DOMAIN 31 355 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 356 372 POTENTIAL.
FT DOMAIN 373 402 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 346 346 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 402 AA; 45211 MW; 78A059332D0238C CRC64;

Query Match 89.7%; Score 26; DB 1; Length 402;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 337 TKPPK 341

RESULT 40
ID_CCKR_HUMAN STANDARD; PRT; 428 AA.
AC P32238;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cholecystokinin type A receptor (CCK-A receptor) (CCK-AR).
GN CCKAR OR CCKRA.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Gall bladder;
RX MEDLINE=9327752; PubMed=8503909;
RA Ulrich C.D., Ferber I., Holicky E., Hadac E., Buell G.,
RA Miller L.J.;
RT "Molecular cloning and functional expression of the human gallbladder
RT cholecystokinin A receptor.",
RL Biochem. Biophys. Res. Commun. 193:204-211(1993).

[2]
RN SEQUENCE FROM N.A.
RP MEDLINE=9343941; PubMed=8343165;
RX Wank S.A., de Weerth A., Pilegna J.R., Huppi K.;
RT "Molecular cloning, functional expression and chromosomal
RT localization of the human cholecystokinin type A receptor.",

Biochem. Biophys. Res. Commun. 194:811-818(1993).

[3]
RN SEQUENCE FROM N.A.
RP MEDLINE=96029343; PubMed=7557108;
RX Miller L.J., Holicky E.L., Ulrich C.D., Wiesen E.D.;
RT "Abnormal processing of the human cholecystokinin receptor gene in
RT association with gallstones and obesity.",
RL Gastroenterology 109:1375-1380(1995).

[4]
RN SEQUENCE FROM N.A.
RP TISSUE=Peripheral blood leukocytes;
RC MEDLINE=20145045; PubMed=10682840;
RX Funakoshi A., Miyasaka K., Matsumoto H., Yamamori S., Takiguchi S.,
RA Katsuka K., Takata Y., Matsue K., Kono A., Shimokata H.;
RT "Gene structure of human cholecystokinin (CCK) type-A receptor: body
RT fat content is related to CCK type-A receptor gene promoter
RT polymorphism.",
RL FEBS Lett. 466:264-266(2000).

[5]
RN SEQUENCE FROM N.A.
RP TISSUE=Stomach;
RC Kopatz S.A., Aronstam R.S., Sharma S.V.;
RT "cDNA clones of human proteins involved in signal transduction
RT cloned by the Guthrie cDNA resource center (www.cdna.org).";
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Receptor for cholecystokinin. Has a 1000 fold affinity
CC for CCK rather than for gastrin. It modulates feeding and
CC dopamine-induced behavior in the central and peripheral nervous
CC system. This receptor mediates its action by association with G
CC proteins that activate a phosphatidylinositol-calcium second
CC messenger system.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; L13605; AAA35659.1; -
DR EMBL; L19315; AAA02819.1; -
DR EMBL; U23430; AAA91123.1; -
DR EMBL; U23427; AAA91123.1; JOINED.
DR EMBL; U23428; AAA91123.1; JOINED.
DR EMBL; U23429; AAA91123.1; JOINED.
DR EMBL; D85606; BAA90879.1; -
DR EMBL; AY322549; AAP84362.1; -
DR PIR; JN0692; JN0692.
DR PDB; 1D6G; 17-NOV-99.
DR PDB; 1H2N; 25-APR-01.
DR Genew; HGNC:1570; CCKAR.
DR MIM; 118444;
DR GO; GO:0008887; C:integral to plasma membrane; TAS.
DR GO; GO:0004951; F:cholecystokinin receptor activity; TAS.
DR GO; GO:0007204; P:cytosolic calcium ion concentration elevation; TAS.
DR GO; GO:0007586; P:digestion; TAS.
DR GO; GO:0007631; P:feeding behavior; TAS.
DR GO; GO:0007584; P:response to nutrients; TAS.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Lipoprotein;
KW Palmitate; 3D-structure.
FT DOMAIN 1 41 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 42 67 1 (POTENTIAL)
FT DOMAIN 68 77 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 78 104 2 (POTENTIAL).
FT DOMAIN 105 115 EXTRACELLULAR (POTENTIAL).

FT TRANSNM 116 137 3 (POTENTIAL).
 FT DOMAIN 138 157 CYTOPLASMIC (POTENTIAL).
 FT TRANSNM 158 178 4 (POTENTIAL).
 FT DOMAIN 179 210 EXTRACELLULAR (POTENTIAL).
 FT TRANSNM 211 234 5 (POTENTIAL).
 FT DOMAIN 235 313 CYTOPLASMIC (POTENTIAL).
 FT TRANSNM 314 334 6 (POTENTIAL).
 FT DOMAIN 335 349 EXTRACELLULAR (POTENTIAL).
 FT TRANSNM 350 373 7 (POTENTIAL).
 FT DOMAIN 374 428 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 10 10 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 114 196 BY SIMILARITY.
 FT LIPID 387 S-palmitoyl cysteine (By similarity).
 SQ SEQUENCE 428 AA; A68FAEDA05E610 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 428;
 Best Local Similarity 80.0%; Pred. No. 3.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
 Db 274 TRPPR 278

RESULT 41
 NOCT_MOUSE STANDARD; PRT; 429 AA.
 AC Q35710; Q9QZG9;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Nocturnin (CCR4 protein homolog).
 GN CCRN4L OR NOC OR CCR4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Brain;
 RA Dupressoir A., Barbot W., Loireau M.P., Heidmann T.;
 RT "Characterization of a mammalian gene related to the yeast CCR4
 RT general transcription factor and revealed by transposon insertion.";
 RL J. Biol. Chem. 274:31068-31075(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Retina;
 RA Wang Y., Osterbur D.L., Megaw P.L., Tosini G., Fukuhara C.,
 RT "Rhythmic expression of Nocturnin mRNA in multiple tissues of the
 RT mouse.";
 RL BMC Dev. Biol. 1:9-9(2001).
 RN [3]
 RP SEQUENCE OF 62-429 FROM N.A.
 RC STRAIN=DBA/2J; TISSUE=Liver;
 RX MEDLINE=97190339; PubMed=9038221;
 RA Pusch A., Dupressoir A., Loireau M.P., Mattei M.-G., Heidmann T.;
 RT "Characterization of two age-induced intracisternal A-particle-related
 RT transcripts in the mouse liver. Transcriptional read-through into an
 RT open reading frame with similarities to the yeast ccr4 transcription
 RT factor.";
 RL J. Biol. Chem. 272:5995-6003(1997).
 CC -!- FUNCTION: Component of the circadian clock or downstream effector
 CC of clock function. Exhibits a high amplitude circadian rhythm with
 CC maximal levels in early evening. In constant darkness or constant
 CC light, the amplitude of the rhythm decreases (by similarity).
 CC -!- SIMILARITY: Belongs to the CCR4/nocturnin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF183960; AAD56547.1; -;
 DR EMBL; AF199491; AAG01384.1; -;
 DR EMBL; U70139; AAB62717.1; ALT_FRAME.
 DR MGD; MG1109382; Ccrn4l.
 DR InterPro; IPR005135; Exo_endo_phos.
 DR Pfam; PF03372; Exo_endo_phos; 1.
 KW Biological rhythms.
 FT CONFLICT 123 125 YOR -> LPA (IN REF. 3).
 SQ SEQUENCE 429 AA; CB9FBE5D84E13942 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 429;
 Best Local Similarity 80.0%; Pred. No. 3.9e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
 Db 118 TRPPR 122

RESULT 42
 NOCT_HUMAN STANDARD; PRT; 431 AA.
 AC Q9UK39; Q9HD94; Q9HD95;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Nocturnin (CCR4 protein homolog).
 GN CCRN4L OR NOC OR CCR4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=99453012; PubMed=10521507;
 RA Dupressoir A., Barbot W., Loireau M.P., Heidmann T.;
 RT "Characterization of a mammalian gene related to the yeast CCR4
 RT general transcription factor and revealed by transposon insertion.";
 RL J. Biol. Chem. 274:31068-31075(1999).
 RN [2]
 RP SEQUENCE OF 67-431 FROM N.A.
 RC Wang Y., Osterbur D.L., Green C.B., Besharse J.C.;
 RT "Mammalian homologs of Xenopus nocturnin: conservation of structure
 RT and circadian regulation.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Component of the circadian clock or downstream effector
 CC of clock function. Exhibits a high amplitude circadian rhythm with
 CC maximal levels in early evening. In constant darkness or constant
 CC light, the amplitude of the rhythm decreases (by similarity).
 CC -!- SIMILARITY: Belongs to the CCR4/nocturnin family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF183961; AAD56548.1; -;
 DR EMBL; AF199492; AAG01387.1; -;
 DR EMBL; AF199493; AAG01388.1; -;
 DR EMBL; AF199494; AAG01389.1; -;
 DR Genew; HGNC:14254; CCRN4L.
 DR GO; GO:0005634; Cinnuleus; TAS.
 DR GO; GO:0003700; F:transcription factor activity; TAS.

DR GO: GO:0006366; P:transcription from Pol II promoter; TAS.
DR InterPro; IPR005135; Exo_endo_phos.
DR Pfam; PF03372; Exo_endo_phos; 1.
KW Biological rhythms. 69 T -> N (IN REF. 2).
FT CONFLICT 77 G -> A (IN REF. 2).
FT CONFLICT 266 A -> T (IN REF. 2); AAG01389).
FT CONFLICT 341 N -> S (IN REF. 2).
SQ SEQUENCE 431 AA; 48150 MW; B61EF484E8D29AF5 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 431;
Best Local Similarity 80.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 120 TRPPR 124

RESULT 43

ENGA_CLOPE STANDARD; PRT; 438 AA.
AC Q8XJK1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE GTP-binding protein engA.
GN ENGA OR CPEI755.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-ester."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -!- FUNCTION: GTPase of unknown physiological role.
CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
CC proteins. EngA subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; AP003191; BAB81461.1; -.
CC HAMAP; MF_00195; 1
DR InterPro; IPR003593; AAA_AtPase.
DR InterPro; IPR005289; GTP-Bindndng_dom.
DR InterPro; IPR006073; GTP1_OBG.
DR InterPro; IPR002917; MMR_HSR1.
DR InterPro; IPR001806; Ras_transfmg.
DR InterPro; IPR005225; Small_GTP.
DR Pfam; PF01926; MMR_HSR1; 1.
DR PRINTS; PRO0326; GTP1_OBG.
DR PRINTS; PRO0449; RASTNSFRMG.
DR SMART; SM00382; AAA; 2.
DR TIGRFAMs; TIGR00650; MG442; 2.
DR TIGRFAMs; TIGR00231; small_GTP; 2.
KW GTP-binding; Repeat; Complete proteome.
FT NP_BIND 10 17 GTP 1 (POTENTIAL).
FT NP_BIND 57 61 GTP 1 (POTENTIAL).
FT NP_BIND 120 123 GTP 1 (POTENTIAL).
FT NP_BIND 183 190 GTP 2 (POTENTIAL).
FT NP_BIND 230 234 GTP 2 (POTENTIAL).

FT NP_BIND 295 298 GTP 2 (POTENTIAL).
SQ SEQUENCE 438 AA; 49586 MW; 9658B0BC98398396 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 438;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 391 TKPPK 395

RESULT 44

SRMB_HAEIN STANDARD; PRT; 439 AA.
AC P44701;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP-dependent RNA helicase srmb homolog.
GN SRMB OR H10422.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKeeney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."

RL Science 269:496-512(1995).
CC -!- FUNCTION: RNA-DEPENDENT ATPASE ACTIVITY. PROBABLY INTERACTS
CC WITH 23S RIBOSOMAL RNA (BY SIMILARITY).
CC -!- SIMILARITY: Belongs to the DEAD box helicase family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; U32725; AAC22078.1; -.
CC PIR; H64066; H64066.
DR HSP; Q58083; LHVS.
DR TIGR; H10422;
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR000629; DEAD_box.
DR InterPro; IPR001650; Helicase_C.
DR Pfam; PF00270; DEAD; 1.
DR Pfam; PF00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00039; DEAD_ATP_HELICASE; 1.
KW Helicase; ATP-binding; RNA-binding; Complete proteome.
FT NP_BIND 48 55 ATP (POTENTIAL).
FT SITE 157 160 DEAD BOX.
SQ SEQUENCE 439 AA; 49805 MW; 7543942CE35B20C4 CRC64;

Query Match 89.7%; Score 26; DB 1; Length 439;
Best Local Similarity 80.0%; Pred. No. 4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY      1 TKPR 5
Db      383 TRPK 387

RESULT 45
VGLD HSVK
ID      VGLD HSVK      STANDARD;      PRT;      442 AA.
AC      P22454;
DT      01-AUG-1991 (Rel. 19, Created)
DT      01-AUG-1991 (Rel. 19, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Glycoprotein D precursor (Glycoprotein 17/18).
GN      GD OR GP17/18 OR 72.
OS      Equine herpesvirus type 1 (strain Kentucky A) (EHV-1).
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC      Alphaherpesvirinae; Varicellovirus.
OX      NCBI_TaxID=10329;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91082407; PubMed=1845821;
RA      Flowers C.C., Eastman E.M., O'Callaghan D.J.;
RT      "Sequence analysis of a glycoprotein D gene homolog within the unique
RT      short segment of the EHV-1 genome.";
RL      Virology 180:175-184(1991).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=92263758; PubMed=1316573;
RA      Colle C.F. III, Flowers C.C., O'Callaghan D.J.;
RT      "Open reading frames encoding a protein kinase, homolog of
RT      glycoprotein gX of pseudorabies virus, and a novel glycoprotein map
RT      within the unique short segment of equine herpesvirus type 1.";
RL      Virology 188:545-557(1992).
CC      -1- SIMILARITY: Belongs to the herpesviruses glycoprotein D family.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M62923; AAA46081.1; -
DR      EMBL; M86931; -; NOT ANNOTATED CDS.
DR      PIR; A38518; VGBSEA.
DR      InterPro; IPR002896; Herpes glycop_D.
DR      Pfam; PF01537; Herpes glycop_D; 1.
KW      Glycoprotein; Signal; Transmembrane.
FT      SIGNAL      1      19      POTENTIAL.
FT      CHAIN      20      442      GLYCOPROTEIN D.
FT      DOMAIN      20      405      EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM      406      422      POTENTIAL.
FT      DOMAIN      423      442      CYTOPLASMIC (POTENTIAL).
FT      CARBOHYD      103      103      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD      111      111      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD      347      347      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD      396      396      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE      442 AA; 49908 MW; 332CDA9C9762F05 CRC64;

Query Match      89.7%; Score 26; DB 1; Length 442;
Best Local Similarity      80.0%; Pred. No. 4e+02;
Matches      4; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

QY      1 TKPR 5
Db      387 TRPK 391

RESULT 46
VGLD HSVB
ID      VGLD HSVB      STANDARD;      PRT;      452 AA.
AC      P24379;
DT      01-MAR-1992 (Rel. 21, Created)
DT      01-MAR-1992 (Rel. 21, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Glycoprotein D precursor (Glycoprotein 17/18).
GN      GD OR GP17/18 OR 72.
OS      Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
OS      Equine herpesvirus type 1 (strain Kentucky D) (EHV-1).
OC      Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC      Alphaherpesvirinae; Varicellovirus.
OX      NCBI_TaxID=31520, 10330;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=AB4P;
RX      MEDLINE=92295566; PubMed=1318606;
RA      Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT      "The DNA sequence of equine herpesvirus-1.";
RL      Virology 189:304-316(1992).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      STRAIN=Kentucky D;
RX      MEDLINE=91108393; PubMed=2177089;
RA      Audonnet J.-C., Winslow J., Allen G., Paoletti E.;
RT      "Equine herpesvirus type 1 unique short fragment encodes
RT      glycoproteins with homology to herpes simplex virus type 1 gD, gI and
RT      gE.";
RL      J. Gen. Virol. 71:2969-2978(1990).
CC      -1- SIMILARITY: Belongs to the herpesviruses glycoprotein D family.
CC      CAUTION: It is uncertain whether Met-1 or Met-51 is the initiator.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; M86664; AAB02507.1; -
DR      PIR; I36802; VGBEG3.
DR      InterPro; IPR002896; Herpes glycop_D.
DR      InterPro; IPR007110; Ig-like.
DR      Pfam; PF01537; Herpes glycop_D; 1.
KW      Glycoprotein; Signal; Transmembrane.
FT      SIGNAL      1      19      POTENTIAL.
FT      CHAIN      20      452      GLYCOPROTEIN D.
FT      DOMAIN      20      405      EXTRACELLULAR (POTENTIAL).
FT      TRANSMEM      406      422      POTENTIAL.
FT      DOMAIN      423      452      CYTOPLASMIC (POTENTIAL).
FT      CARBOHYD      103      103      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD      111      111      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD      347      347      N-LINKED (GLCNAC. . .) (POTENTIAL).
FT      CARBOHYD      396      396      N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ      SEQUENCE      452 AA; 51099 MW; CF51E914F7E29DC CRC64;

Query Match      89.7%; Score 26; DB 1; Length 452;
Best Local Similarity      80.0%; Pred. No. 4.1e+02;
Matches      4; Conservative      1; Mismatches      0; Indels      0; Gaps      0;

QY      1 TKPR 5
Db      387 TRPK 391

RESULT 47
ENGA-BIFLO
ID      ENGA-BIFLO      STANDARD;      PRT;      463 AA.
AC      Q8G6J8;
DT      15-MAR-2004 (Rel. 43, Created)
DT      15-MAR-2004 (Rel. 43, Last sequence update)
DT      15-MAR-2004 (Rel. 43, Last annotation update)

```

DE GTP-binding protein engA.
GN ENGA OR BL0738.
OS Bifidobacterium longum.
CC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
CC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]_TaxID=216816;
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22994977; PubMed=13381787;
RA Scheil M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA Pessi G., Zwaan M.-C., Desiere F., Bork P., Delley M.,
RA Pridmore R.D., Arigoni F.,
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
to the human gastrointestinal tract".
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
CC -!- FUNCTION: GTPase of unknown physiological role.
CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
CC proteins. EngA subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; AEO14696; AAN24555.1; -.
CC HAMAP; MF_00195; 1.
CC InterPro; IPR003593; AAA ATPase.
CC InterPro; IPR005289; GTP-binding_dom.
CC InterPro; IPR006073; GTP1_OBG.
CC InterPro; IPR002917; MMR_HSR1.
CC InterPro; IPR005225; Small_GTP.
CC Pfam; PF01926; MMR_HSR1; 1.
CC PRINTS; PR00326; GTP1_OBG.
CC SMART; SMO0382; AAA; 2.
CC TIGRFAMs; TIGR00650; MG442; 2.
CC TIGRFAMs; TIGR00231; small_GTP; 2.
CC TIGRFAMs; Repeat; Complete proteome.
CC NP_BIND 33 40 GTP 1 (POTENTIAL).
CC NP_BIND 80 84 GTP 1 (POTENTIAL).
CC NP_BIND 142 145 GTP 1 (POTENTIAL).
CC NP_BIND 208 215 GTP 2 (POTENTIAL).
CC NP_BIND 255 259 GTP 2 (POTENTIAL).
CC NP_BIND 320 323 GTP 2 (POTENTIAL).
CC NP_BIND 463 AA; 51186 MW; 7F986F8E041367AE CRC64;
SQ SEQUENCE 463 AA; 51186 MW; 7F986F8E041367AE CRC64;
Query Match 89.7%; Score 26; DB 1; Length 463;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Cy 1 TKPFR 5
Db 414 TRPFR 418
RESULT 48
EM55 HUMAN STANDARD; PRT; 466 AA.
AC Q00013;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 55 kDa erythrocyte membrane protein (p55) (Membrane protein,
DE palmitoylated 1).
GN MPPI OR EMP55.
OS Homo sapiens (Human).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RC TISSUE=Reticulocytes;
RX MEDLINE=91319732; PubMed=1713685;
RA Ruff P., Speicher D.W., Husain-Chishti A.;
RT "Molecular identification of a major palmitoylated erythrocyte
RT membrane protein containing the src homology 3 motif".
RL Proc. Natl. Acad. Sci. U.S.A. 88:6595-6599(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93244792; PubMed=1301163;
RA Metzberg A.B., Gitschler J.;
RT "The gene encoding the palmitoylated erythrocyte membrane protein,
RT p55, originates at the CpG island 3' to the factor VIII gene".
RL Hum. Mol. Genet. 1:97-101(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares W.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullay S.J.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., McKernan K.J., Malek J.A., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences".
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBUNIT: Interacts with DLG5.
CC -!- SUBCELLULAR LOCATION: Membrane-associated.
CC -!- PTM: Extensively palmitoylated.
CC -!- SIMILARITY: Belongs to the MAGUK family.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- SIMILARITY: Contains 1 guanylate kinase-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; M64925; AAA60059.1; -.
CC ENBL; M87059; AAA60060.1; -.
CC ENBL; U39611; AAD14835.1; -.
CC ENBL; BC002392; AAH02392.1; -.
CC HSP; O14936; 1KWA.
CC Genew; HGNC:7219; MPPI.
CC MIM; 305360; -.
CC GO; GO:0003887; C:integral to plasma membrane; TAS.
CC GO; GO:0005624; C:membrane fraction; TAS.
CC GO; GO:0004385; F:guanylate kinase activity; TAS.
CC GO; GO:0007165; P:signal transduction; TAS.
CC InterPro; IPR008144; Guanylate_kin.
CC InterPro; IPR008145; Guanylt/Gs.
CC InterPro; IPR001478; PDZ.
CC InterPro; IPR001452; SH3.
CC Pfam; PF00625; Guanylate_kin; 1.
CC Pfam; PF00595; PDZ; 1.
CC Pfam; PF00018; SH3; 1.
CC ProDom; PD000066; SH3; 1.


```

DR SMART, SM00072; GUKC; 1.
DR SMART, SM00228; PDZ; 1.
DR SMART, SM00326; SH3; 1.
DR PROSITE; PS00856; GUANYLATE_KINASE_1; 1.
DR PROSITE; PS00052; GUANYLATE_KINASE_2; 1.
DR PROSITE; PS0106; PDZ; 1.
DR PROSITE; PS00002; SH3; 1.
KW Membrane, Erythrocyte, SH3 domain; Lipoprotein; Palmitate;
FT DOMAIN 71 152 PDZ.
FT DOMAIN 158 228 SH3.
FT DOMAIN 280 466 GUANYLATE KINASE.
FT VARIANT 448 448 E -> Q (in dbSNP:14092).
FT FTID=VAR_011914.
SQ SEQUENCE 466 AA; 52296 MW; DC68AA68F49A26E CRC64;

Query Match 89.7%; Score 26; DB 1; Length 466;
Best Local Similarity 80.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 318 TRPPR 322

RESULT 49
ENP6_HUMAN STANDARD; PRT; 484 AA.
ID ENP6_HUMAN STANDARD; PRT; 484 AA.
AC O75354; Q9UUD1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 6 (EC 3.6.1.16)
DE (NTPase6) (CD39 antigen-like 2).
GN ENTPD6 OR CD39L2.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Keratinocytes;
RX MEDLINE=98341119; PubMed=9676430;
RA Chadwick B.P., Frischauf A.-M.;
RT "The CD39-like gene family: identification of three new human members
RT (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
RT the gene family from Drosophila melanogaster.";
RL Genomics 50:357-367(1998).
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Graiham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Leivaesalho M.H., Leverha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConnachie L.J., McLeay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Pratchalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen R., Sims S.,
RA Stuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,

```

```

RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
CC -1- FUNCTION: Might support glycosylation reactions in the Golgi
CC apparatus and when released from cells, might catalyze the
CC hydrolysis of extracellular nucleotides. Hydrolyzes preferentially
CC nucleoside 5'-diphosphates, nucleoside 5'-triphosphates are
CC hydrolyzed only to a minor extent, there is no hydrolysis of
CC nucleoside 5'-monophosphates. The order of activity with different
CC substrates is GDP > IDP > UDP = CDP >> ADP (By similarity).
CC -1- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -1- COFACTOR: Requires calcium and magnesium (By similarity).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI. BUT ALSO
CC OCCURS IN A SOLUBLE EXTRACELLULAR FORM (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: Expressed in most tissues.
CC -1- SIMILARITY: Belongs to the GDAL / CD39 NTPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF039916; AAC39883.1; -
CC EMBL; AL035252; CAB41571.1; -
CC Genew; HGNC:3368; ENTPD6.
CC MIM; 603160; -
CC InterPro; IPR000407; GDAL_CD39_NTPase.
CC Pfam; PF01150; GDAL_CD39_1.
CC PROSITE; PS01238; GDAL_CD39_NTPase; FALSE NEG.
KW Hydrolyase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW Signal-anchor; Go-gi stack.
FT DOMAIN 1 39
FT TRANSMEM 40 60
CYTOPLASMIC (POTENTIAL).
SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
LUMINAL (POTENTIAL).
FT DOMAIN 61 484
FT CARBOHYD 220 220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 138 138 V -> L (IN REF. 2).
FT CONFLICT 202 202 E -> K (IN REF. 2).
SQ SEQUENCE 484 AA; 53233 MW; 27334E29DDB8D64C CRC64;

Query Match 89.7%; Score 26; DB 1; Length 484;
Best Local Similarity 80.0%; Pred. No. 4.4e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 121 TRPPR 125

RESULT 50
AMPA_AERPE STANDARD; PRT; 524 AA.
ID AMPA_AERPE STANDARD; PRT; 524 AA.
AC Q9Y935;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
DE (LAP) (Leucyl aminopeptidase).
GN PEPA OR APE2450.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;

```

PX MEDLINE-99310339; PubMed-10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Hatakeyama Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
CC -!- FUNCTION: Presumably involved in the processing and regular
CC turnover of intracellular proteins. Catalyzes the removal of
CC unsubstituted N-terminal amino acids from various peptides (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-|-
CC including Pro although not Arg or Lys, and Xbb may be Pro.
CC -!- COFACTOR: Binds 2 manganese ions per subunit (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to peptidase family M17.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000064; BA81465.1; -;
DR PIR: A72476; A72476.
DR HSPF: P00727; ILAP.
DR MEROPS: M17.0PW; -;
DR HAMAP: MF 00181; -; 1.
DR InterPro: IPR000819; Peptidase M17_C.
DR InterPro: IPR008283; Peptidase M17_N.
DR Pfam: PF00883; Peptidase M17; 1.
DR Pfam: PF02789; Peptidase M17_N; 1.
DR PRINTS: PR0848; LAMNOPPTDASE.
DR PROSITE: PS00631; CYTOSOL_AP; 1.
KW Hydrolyase; Aminopeptidase; Manganese; Complete proteome.
FT METAL 280 280 MANGANESE 2 (BY SIMILARITY).
FT METAL 285 285 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 303 303 MANGANESE 2 (BY SIMILARITY).
FT METAL 362 362 MANGANESE 1 (BY SIMILARITY).
FT METAL 364 364 MANGANESE 1 AND 2 (BY SIMILARITY).
FT ACT_SITE 292 292 POTENTIAL.
FT ACT_SITE 366 366 POTENTIAL.
SQ SEQUENCE 524 AA; 55609 MW; 0D6161A4D814C75D CRC64;

Query Match 89.7%; Score 26; DB 1; Length 524;
Best Local Similarity 80.0%; Pred. No. 4.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPR 5
|:|
Db 37 TRPR 41

Search completed: March 3, 2004, 12:17:40
Job time : 14 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 12:15:23 ; Search time 20 Seconds
(without alignments)
24.048 Million cell updates/sec

Title: US-09-871-974-2

Perfect score: 29

Sequence: 1 TKPPR 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

PIR 78:*

1: Piri:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	29	100.0	142	2 C82829	conserved hypothet
2	29	100.0	147	2 AH3157	conserved hypothet
3	29	100.0	191	2 T47369	hypothetical prote
4	29	100.0	201	2 S76026	hypothetical prote
5	29	100.0	222	2 G98129	hypothetical prote
6	29	100.0	226	2 T25360	hypothetical prote
7	29	100.0	278	2 S75883	hypothetical prote
8	29	100.0	283	2 T18299	hypothetical prote
9	29	100.0	350	2 AE3171	ATP-dependent DNA
10	29	100.0	351	2 T29369	hypothetical prote
11	29	100.0	370	2 AC0173	probable iron-sulf
12	29	100.0	375	2 D84715	probable protein k
13	29	100.0	410	2 S38238	hypothetical prote
14	29	100.0	415	2 F96499	hypothetical prote
15	29	100.0	415	2 C84829	hypothetical prote
16	29	100.0	442	2 S61165	hypothetical prote
17	29	100.0	493	2 S73752	hypothetical prote
18	29	100.0	535	2 T21775	hypothetical prote
19	29	100.0	604	2 S56027	hypothetical prote
20	29	100.0	615	2 JC7576	transcription fact
21	29	100.0	619	2 A43361	Sts-related transc
22	29	100.0	722	2 T22359	hypothetical prote
23	29	100.0	816	2 T45684	hypothetical prote
24	29	100.0	919	2 A41275	DNA ligase (ATP) (
25	29	100.0	941	2 T51135	ligand-gated chann
26	29	100.0	962	2 D86186	hypothetical prote
27	29	100.0	976	2 T51137	ionotropic glutama
28	29	100.0	1235	1 FWBYH	potassium transpor
29	29	100.0	1241	2 JU0466	potassium transpor

hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
vacuolar precursor
ribosomal protein
hypothetical prote
hypothetical prote
hypothetical prote
transposase homolo
conserved homolo
hypothetical prote
hypothetical prote
acidic endoprotein
major core protein
hypothetical prote
guanylate kinase
guanylate kinase
guanylate kinase N
AcOrf-132 protein
ACMPV orf132 - Bo
hypothetical prote
hypothetical prote
tRNA pseudouridine
hypothetical prote
spermidine synthas
hypothetical prote
probable dihydrodi
hypothetical prote
putative transcrip
LuxA-related prote
calcium-binding pr
genome polyprotein
hypothetical prote
ATP-dependent DNA
site-specific DNA
type II DNA modifi
hypothetical prote
brefeldin A estera
tRNA (5-methylamin
probable GTP bindi
cholecystokinin ty
hypothetical prote
ATP-dependent RNA
Glycoprotein D pre
Glycoprotein D pre
55K erythrocyte me
homeotic protein H
probable cytosol a
phosphoinositide-s
oligopeptide trans
gag polyprotein -
hypothetical prote
phosphoinositide-s
phosphoinositide-s
phosphoinositide-s
phosphoinositide-s
phosphoinositide-s
phospholipase C1
phosphoinositide-s
phosphoinositide-s
phosphoinositide-s


```

Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
    |||||
Db 208 TKPPR 212

RESULT 9
AE3171
ATP-dependent DNA ligase Atu5097 [imported] - Agrobacterium tumefaciens (strain C58, Dup
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C/Accession: AE3171
R/Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woc, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AE3171
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-350 <KUR>
A/Cross-references: GB:AB008687; PIDN:AAL45787.1; PID:gl7743523; GSPDB:GN00188
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu5097
A/Genome: plasmid

Query Match 100.0%; Score 29; DB 2; Length 350;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
    |||||
Db 2 TKPPR 6

RESULT 10
T29369
Hypothetical protein ZC404.8 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T29369
R/Bentley, D.; Le, T.T.
submitted to the EMBL Data Library, April 1996
A/Description: The sequence of C. elegans cosmid ZC404.
A/Reference number: Z20614
A/Accession: T29369
A/Status: preliminary; translated from GB/EMBL/DD5J
A/Molecule type: DNA
A/Residues: 1-351 <BEN>
A/Cross-references: EMBL:U5363; PIDN:AAA97963.1; GSPDB:GN00023; CESP:ZC404.8
A/Experimental source: strain Bristol N2; Clone ZC404
C/Genetics:
A/Gene: CESP:ZC404.8
A/Map position: 5
A/Introns: 17/2; 52/2; 73/2; 312/1

Query Match 100.0%; Score 29; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
    |||||
Db 199 TKPPR 203

RESULT 11
AC0173
probable iron-sulfur binding protein YP01417 [imported] - Yersinia pestis (strain CO92
C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C/Accession: AC0173
R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell
Nature 413, 523-527, 2001
A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A/Reference number: AB0001; MUID:21470413; PMID:11586360
A/Accession: AC0173
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-370 <KUR>
A/Cross-references: GB:AL590842; PIDN:CAC90246.1; PID:gl5979466; GSPDB:GN00175
C/Genetics:
A/Gene: YP01417

Query Match 100.0%; Score 29; DB 2; Length 370;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
    |||||
Db 266 TKPPR 270

RESULT 12
D84715
Probable protein kinase [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: D84715
R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617157
A/Accession: D84715
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-375 <STO>
A/Cross-references: GB:AE002093; NID:g3201626; PIDN:AAC20735.1; GSPDB:GN00139
C/Genetics:
A/Gene: At2g31010
A/Map position: 2

Query Match 100.0%; Score 29; DB 2; Length 375;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
    |||||
Db 121 TKPPR 125

RESULT 13
S38238
Hypothetical protein - Coxiella burnetii
C/Species: Coxiella burnetii
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 08-Oct-1999
C/Accession: S38238
R/Thiele, D.; Willems, H.; Haas, M.; Krause, H.
submitted to the EMBL Data Library, October 1993
A/Reference number: S38215
A/Accession: S38238
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-410 <THI>
A/Cross-references: EMBL:X75356; NID:g407370; PIDN:CAA53126.1; PID:g407394

Query Match 100.0%; Score 29; DB 2; Length 410;

```

```

Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 102 TKPPR 106

RESULT 14
F96499
hypothetical protein T10P12.9 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
R;Accession: F96499
R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: F96499
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-415 <STO>
A;Cross-references: GB:AE005173; NID:G5080765; PIDN:AA039275.1; GSPDB:GN00141
C;Genetics:
A;Gene: T10P12.9
A;Map position: 1

Query Match 100.0%; Score 29; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 309 TKPPR 313

RESULT 15
C84829
hypothetical protein At2g40420 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84829
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: C84829
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-415 <STO>
A;Cross-references: GB:AE002093; NID:G6598346; PIDN:AAB87575.2; GSPDB:GN00139
C;Genetics:
A;Gene: At2g40420
A;Map position: 2

Query Match 100.0%; Score 29; DB 2; Length 415;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 169 TKPPR 173

RESULT 16
S61165
hypothetical protein YDR370c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein D9481.14
C;Species: Saccharomyces cerevisiae
C;Date: 23-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 19-Apr-2002
C;Accession: S61165
R;Ding, H.
submitted to the EMBL Data Library, June 1995
A;Description: The sequence of S. cerevisiae cosmid 9481.
A;Reference number: S61159
A;Accession: S61165
A;Molecule type: DNA
A;Residues: 1-442 <DIN>
A;Cross-references: EMBL:U29373; NID:G649184; PID:G849191; GSPDB:GN00004; MIPS:YDR370c
A;Experimental source: strain S288C (AB972)
C;Genetics:
A;Gene: MIPS:YDR370c
A;Cross-references: SGD:S0002778
A;Map position: 4R
C;Superfamily: Saccharomyces cerevisiae hypothetical protein YDR370c

Query Match 100.0%; Score 29; DB 2; Length 442;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 81 TKPPR 85

RESULT 17
S73752
hypothetical protein A05_orf493 - Mycoplasma pneumoniae (strain ATCC 29342)
C;Species: Mycoplasma pneumoniae
A;Variety: ATCC 29342
C;Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 15-Sep-2000
C;Accession: S73752
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae
A;Reference number: S73327; MUID:97105885; PMID:8948633
A;Accession: S73752
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-493 <HIM>
A;Cross-references: EMBL:AE000041; GB:U00089; NID:G1674104; PIDN:AAB96074.1; PID:G16741
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C;Genetics:
A;Genetic code: SGCG
C;Superfamily: Mycoplasma pneumoniae hypothetical protein H08_orf445

Query Match 100.0%; Score 29; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 462 TKPPR 466

RESULT 18
T21775
hypothetical protein F35E2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21775
R;Lennard, N.
submitted to the EMBL Data Library, November 1996
A;Reference number: Z19471
A;Accession: T21775
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA

```


A;Residues: 1-575 <WIL>
 A;Cross-references: EMBL:281528; PIDN:CAB04281.1; GSPDB:GN00019; CESP:F35E2.5
 A;Experimental source: clone F35E2
 C;Genetics:
 A;Gene: CESP:F35E2.5
 A;Map position: 1
 A;Introns: 44/1; 109/2; 209/3; 250/1; 367/1; 413/1; 480/3; 508/3

Query Match 100.0%; Score 29; DB 2; Length 575;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPRP 5
 |||||
 Db 284 TKPRP 288

RESULT 19

S56027
 hypothetical protein YUL083w - yeast (Saccharomyces cerevisiae)
 N;Alternate names: hypothetical protein YJ002
 C;Species: Saccharomyces cerevisiae
 C;Date: 05-May-1995 #sequence_revision 08-Sep-1995 #text_change 19-Apr-2002
 C;Accession: S56027; S56850; S57742
 R;Mitsuga, T.; Schaaff-Gerstenschlaeger, I.; Chaiwatwis, N.; Baur, A.; Boles, E.; Fournie
 Yeast 11, 681-689, 1995
 A;Title: Sequence analysis of a 33.1 kb fragment from the left arm of Saccharomyces cere
 ter domain and a putative alpha-2-SCB-alpha-2 binding site.
 A;Reference number: S56016; MUID:96093911; PMID:7483841
 A;Accession: S56027
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-604 <MIO>
 A;Cross-references: EMBL:X83502; NID:G929861; PIDN:CAB58487.1; PID:G929873
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, December 1994
 R;Mitsuga, T.; Schaaff-Gerstenschlaeger, I.; Baur, A.; Boles, E.; Chaiwatwis, N.; Fournie
 submitted to the Protein Sequence database, September 1995
 A;Reference number: S56855
 A;Accession: S56850
 A;Molecule type: DNA
 A;Residues: 1-604 <MIW>
 A;Cross-references: EMBL:249358; NID:G1008247; PIDN:CAB89376.1; PID:G1008248; MIPS:YJL08
 R;Sor, F.J.
 Submitted to the EMBL Data Library, June 1995
 A;Reference number: S57731
 A;Accession: S57742
 A;Molecule type: DNA
 A;Residues: 1-604 <SOR>
 A;Cross-references: EMBL:X88851; NID:G895892; PIDN:CAB61318.1; PID:G895904
 C;Genetics:
 A;Cross-references: SGD:S0003619
 A;Map position: 10L

Query Match 100.0%; Score 29; DB 2; Length 604;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPRP 5
 |||||
 Db 151 TKPRP 155

RESULT 20

JC7576
 transcription factor Elf-1, type 1 - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
 C;Accession: JC7576
 R;Nishiyama, C.; Takahashi, K.; Nishiyama, M.; Okumura, K.; Ra, C.; Ohtake, Y.; Yokota,
 Biosci. Biotechnol. Biochem. 64, 2601-2607, 2000
 A;Title: Splice isoforms of transcription factor Elf-1 affecting its regulatory function
 A;Reference number: JC7576; MUID: 21077473; PMID:11210123
 A;Contents: Mast cell line, RBL-2H3

A;Accession: JC7576
 A;Molecule type: mRNA
 A;Residues: 1-615 <NIS>
 A;Cross-references: DBJ:AB030215
 C;Comment: This protein, as a key transcription factor for immune-related genes, has tr
 C;Genetics:
 A;Gene: Elf-1

Query Match 100.0%; Score 29; DB 2; Length 615;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPRP 5
 |||||
 Db 179 TKPRP 183

RESULT 21

A43361
 Ets-related transcription factor Elf-1 - human
 N;Alternate names: E74-like factor Elf-1
 C;Species: Homo sapiens (man)
 C;Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 07-May-1999
 C;Accession: A43361; A42122
 R;Laiden, J.M.; Wang, C.Y.; Petryniak, B.; Markovitz, D.M.; Nabel, G.J.; Thompson, C.B
 J. Virol. 66, 5890-5897, 1992
 A;Title: A novel Ets-related transcription factor, Elf-1, binds to human immunodeficien
 A;Reference number: A43361; MUID:92407982; PMID:1527846
 A;Accession: A43361
 A;Molecule type: mRNA
 A;Residues: 1-619 <LEI>
 A;Cross-references: GB:M82882
 R;Thompson, C.B.; Wang, C.Y.; Ho, I.C.; Bohjanen, P.R.; Petryniak, B.; June, C.H.; Mie
 Mol. Cell. Biol. 12, 1043-1053, 1992
 A;Title: cis-acting sequences required for inducible interleukin-2 enhancer function b
 A;Reference number: A42122; MUID:92188836; PMID:1545787
 A;Accession: A42122
 A;Status: preliminary; not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 204-282, 'G', 284-289 <THO>
 A;Experimental source: T-cells
 A;Note: sequence extracted from NCBI backbone (NCBI:P:88288)
 C;Genetics:
 A;Gene: GDB:ELF1
 A;Cross-references: GDB:131648
 A;Map position: lp36-lp36
 C;Superfamily: ets DNA-binding domain homology
 C;Keywords: DNA binding; transcription regulation
 F:210-290/Domain: ets DNA-binding domain homology <ETS>

Query Match 100.0%; Score 29; DB 2; Length 619;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPRP 5
 |||||
 Db 180 TKPRP 184

RESULT 22

T22359
 hypothetical protein F47G4.4 - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T22359
 R;White, S.
 submitted to the EMBL Data Library, September 1997
 A;Reference number: Z19553
 A;Accession: T22359
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-722 <WIL>
 A;Cross-references: EMBL:Z99171; PIDN:CAB16313.1; GSPDB:GN00019; CESP:F47G4.4

A;Experimental source: clone F47G4

C;Genetics:

A;Gene: CBSP.F47G4.4

A;Map position: 1

A;Introns: 571; 43/3; 129/3; 176/3; 226/1; 259/3; 319/3; 365/3; 465/2; 576/1; 631/1; 688

Query Match 100.0%; Score 29; DB 2; Length 722;

Best Local Similarity 100.0%; Pred. No. 3.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

DB 402 TKPPR 406

RESULT 23

T45684

hypothetical protein F14P22.230 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 04-Mar-2000

C;Accession: T45684

R;D'Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Lemcke, K.

submitted to the Protein Sequence Database, January 2000

A;Reference number: 223011

A;Accession: T45684

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-816 <DAN>

A;Cross-references: EMBL:AL137082

A;Experimental source: cultivar Columbia; BAC clone F14P22

C;Genetics:

A;Map position: 3

A;Introns: 83/3; 151/3; 209/3; 296/2; 336/2; 366/2; 398/2; 429/2; 469/3; 562/1; 603/2; 6

A;Note: F14P22.230

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo

Query Match

Best Local Similarity 100.0%; Score 29; DB 2; Length 816;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

DB 625 TKPPR 629

RESULT 24

A41275

DNA ligase (ATP) (EC 6.5.1.1) I - human

N;Alternate names: DNA joinase; DNA repair enzyme; polydeoxyribonucleotide synthase (ATP

C;Species: Homo sapiens (man)

C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 03-Feb-2003

C;Accession: A36048; A41275

R;Barnes, D.E.; Johnston, L.H.; Kodama, K.; Tomkinson, A.E.; Lasko, D.D.; Lindahl, T.

Proc. Natl. Acad. Sci. U.S.A. 87, 6679-6683, 1990

A;Title: Human DNA ligase I cDNA: cloning and functional expression in Saccharomyces cer

A;Reference number: A36048; MUID:90370849; PMID:2204063

A;Accession: A36048

A;Molecule type: mRNA

A;Residues: 1-919 <BAR>

A;Cross-references: GB:M36067; NID:q187142; PIDN:AAA59518.1; PID:g187143

R;Petrini, J.H.J.; Huwiler, K.G.; Weaver, D.T.

Proc. Natl. Acad. Sci. U.S.A. 88, 7615-7619, 1991

A;Title: A wild-type DNA ligase I gene is expressed in Bloom's syndrome cells.

A;Reference number: A41275; MUID:91352039; PMID:1881902

A;Accession: A41275

A;Molecule type: mRNA

A;Residues: 716-753 <PER>

C;Genetics:

A;Gene: GDB:LIG1

A;Cross-references: GDB:127274; OMIM:126391

A;Map position: 19q13.3-19q13.3

C;Keywords: DNA repair; ligase; phosphoprotein

F;568/Active site: Lys (covalent AMP-binding) #status predicted

Query Match 100.0%; Score 29; DB 2; Length 919;

Best Local Similarity 100.0%; Pred. No. 4.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

DB 218 TKPPR 222

RESULT 25

T51135

ligand-gated channel-like protein precursor [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 28-Jul-2000

C;Accession: T51135

R;Meyerhoff, O.; Hedrich, R.; Becker, D.

submitted to the EMBL Data Library, July 1999

A;Description: Characterization of ligand-gated channel-like proteins in higher plants

A;Reference number: Z25308

A;Accession: T51135

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-941 <MEY>

A;Cross-references: EMBL:AF167355; PIDN:AAD47833.1

A;Experimental source: cultivar Columbia

C;Genetics:

A;Gene: GLUR3

A;Map position: 1

Query Match

Best Local Similarity 100.0%; Score 29; DB 2; Length 941;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

DB 468 TKPPR 472

RESULT 26

D86186

hypothetical protein [imported] - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C;Accession: D86186

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.

ansen, N.P.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lucos, J.S.; Maiti, R.; Marziani

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: D86186

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-962 <STO>

A;Cross-references: GB:AE005172; NID:92388577; PIDN:AAB71458.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

Query Match

Best Local Similarity 100.0%; Score 29; DB 2; Length 962;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

DB 468 TKPPR 472

RESULT 27

TS1137
 Ionotropic glutamate receptor homolog GLR4 [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 28-Jul-2000 #sequence_revision 28-Jul-2000 #text_change 28-Jul-2000
 C/Accession: TS1137
 R/Davenport, R.J.; Kiegle, E.A.; Tester, M.
 submitted to the EMBL Data Library, September 1999
 A/Description: Cloning of an ionotropic glutamate receptor homolog from Arabidopsis thaliana
 A/Reference number: Z5310
 A/Accession: TS1137
 A/Status: preliminary; translated from GB/EMBL/DDRJ
 A/Molecule type: mRNA
 A/Residues: 1-976 <DAV>
 A/Cross-references: EMBL:AF183932; PIDN:AAF01294.1
 C/Genetics:
 A/Gene: GLR4

Query Match 100.0%; Score 29; DB 2; Length 976;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
 |||||
 Db 468 TKPPR 472

RESULT 28

FWBKH
 potassium transport protein TRK1, high-affinity - yeast (Saccharomyces cerevisiae)
 N/Alternate names: protein J0693; protein YJL129c
 C/Species: Saccharomyces cerevisiae
 C/Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 27-Oct-2003
 C/Accession: S05849; S56910
 R/Gaber, R.F.; Styles, C.A.; Fink, G.R.
 Mol. Cell. Biol. 8, 2848-2859, 1988
 A/Title: TRK1 encodes a plasma membrane protein required for high-affinity potassium transport
 A/Reference number: S05849; MUID:86302204; PMID:3043197
 A/Accession: S05849
 A/Molecule type: DNA
 A/Residues: 1-1235 <GAB>
 A/Cross-references: EMBL:M21328; NID:gl71803; PIDN:AAA34728.1; PID:gl71804
 R/Ciepluch, C.; Kordes, E.; Pujol, A.; Jauniaux, J.C.
 submitted to the Protein Sequence Database, September 1995
 A/Reference number: S56891
 A/Accession: S56910
 A/Molecule type: DNA
 A/Residues: 1-1235 <CZI>
 A/Cross-references: EMBL:Z49404; NID:gl008329; PIDN:CAA89424.1; PID:gl008330; GSPDB:GN000000000

A/Gene: SGD:TRK1; MIPS:YJL129c
 A/Cross-references: SGD:S0003665; MIPS:YJL129c

C/Superfamily: potassium transport protein TRK1/TRK2
 C/Keywords: ion transport; potassium transport; transmembrane protein
 F/50-70/Domain: transmembrane #status predicted <TM01>
 F/72-96/Domain: transmembrane #status predicted <TM02>
 F/107-127/Domain: transmembrane #status predicted <TM03>
 F/778-800/Domain: transmembrane #status predicted <TM04>
 F/812-834/Domain: transmembrane #status predicted <TM05>
 F/839-859/Domain: transmembrane #status predicted <TM06>
 F/862-882/Domain: transmembrane #status predicted <TM07>
 F/898-918/Domain: transmembrane #status predicted <TM08>
 F/923-943/Domain: transmembrane #status predicted <TM09>
 F/972-991/Domain: transmembrane #status predicted <TM10>
 F/1079-1096/Domain: transmembrane #status predicted <TM11>
 F/1111-1132/Domain: transmembrane #status predicted <TM12>
 F/738/Binding site: Arg (Lys) #status predicted

Query Match 100.0%; Score 29; DB 1; Length 1235;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
 |||||
 Db 438 TKPPR 442

RESULT 29

JU0466
 potassium transport protein TRK1, high-affinity - yeast (Saccharomyces cerevisiae) (str)
 N/Alternate names: protein YJL129c
 C/Species: Saccharomyces cerevisiae
 A/Variety: Saccharomyces uvarum
 C/Date: 13-Jul-1990 #sequence_revision 28-Oct-1994 #text_change 27-Oct-2003
 C/Accession: JU0466
 R/Anderson, J.A.; Best, L.A.; Gaber, R.F.
 Gene 99, 39-46, 1991
 A/Title: Structural and functional conservation between the high-affinity K⁺ transporter
 A/Reference number: JU0466; MUID:91216443; PMID:2022322
 A/Accession: JU0466
 A/Molecule type: DNA
 A/Residues: 1-1241 <AND>
 A/Cross-references: GB:M57508; NID:gl71640; PIDN:AAA34661.1; PID:gl71641
 A/Note: the source is designated as Saccharomyces uvarum
 C/Genetics:
 A/Gene: SGD:TRK1
 A/Cross-references: SGD:S0003665; MIPS:YJL129c

A/Map Position: 10L
 C/Superfamily: potassium transport protein TRK1/TRK2
 C/Keywords: ion transport; potassium transport; transmembrane protein
 F/50-70/Domain: transmembrane #status predicted <TM01>
 F/78-98/Domain: transmembrane #status predicted <TM02>
 F/107-117/Domain: transmembrane #status predicted <TM03>
 F/784-806/Domain: transmembrane #status predicted <TM04>
 F/819-840/Domain: transmembrane #status predicted <TM05>
 F/844-864/Domain: transmembrane #status predicted <TM06>
 F/868-888/Domain: transmembrane #status predicted <TM07>
 F/904-924/Domain: transmembrane #status predicted <TM08>
 F/929-949/Domain: transmembrane #status predicted <TM09>
 F/1084-1104/Domain: transmembrane #status predicted <TM10>
 F/1117-1137/Domain: transmembrane #status predicted <TM11>

Query Match 100.0%; Score 29; DB 2; Length 1241;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
 |||||
 Db 444 TKPPR 448

RESULT 30

G81008
 hypothetical protein NMB2072 [imported] - Neisseria meningitidis (strain MC58 serogroup)
 C/Species: Neisseria meningitidis
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C/Accession: G81008
 R/Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A/Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
 A/Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A/Reference number: AB1000; MUID:20175755; PMID:10710307
 A/Accession: G81008
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-47 <TET>

A/Cross-references: GB:AB002557; GB:AB002098; NID:g7227332; PIDN:AAF42391.1; PID:g7227
 A/Experimental source: serogroup B, strain MC58
 C/Genetics:
 A/Gene: NMB2072

Query Match 89.7%; Score 26; DB 2; Length 47;
 Best Local Similarity 80.0%; Pred. No. 85;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
|:|:|
20 TRPPR 24

Db

RESULT 31

hypothetical protein aal2639 [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002

C;Accession: AH2135

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AH2135

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-50 <KUR>

A;Cross-references: GB:BA000019; PIDN:BA074338.1; PID:g17131732; GSPDB:GN00179

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: aal2639

Query Match 89.7%; Score 26; DB 2; Length 50;

Best Local Similarity 80.0%; Pred. No. 91;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
|:|:|
20 TKPPK 24

Db

RESULT 32

hypothetical protein NMB0921 [imported] - Neisseria meningitidis (strain MC58 serogroup C)

C;Species: Neisseria meningitidis

C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001

C;Accession: C81143

R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M. Science 287, 1809-1815, 2000

A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Venter, A.; Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.

A;Reference number: AB1000; MUID:20175755; PMID:10710307

A;Accession: C81143

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-66 <TET>

A;Cross-references: GB:AE002443; GB:AE002098; NID:g7226149; PIDN:AAF41329.1; PID:g722615

A;Experimental source: serogroup B, strain MC58

C;Genetics:

A;Gene: NMB0921

Query Match 89.7%; Score 26; DB 2; Length 66;

Best Local Similarity 80.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
|:|:|
39 TKPPK 43

Db

RESULT 33

hypothetical protein SC128.04c - Streptomyces coelicolor

C;Species: Streptomyces coelicolor

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C;Accession: T36774

R;Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the ENEL Data Library, July 1999

A;Reference number: Z21574

A;Accession: T36774

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-68 <SEE>

A;Cross-references: EMBL:AL096844; PIDN:CAB50878.1; GSPDB:GN00070; SCORDB:SCI28.04c

A;Experimental source: strain A3(2)

C;Genetics:

A;Gene: SCORDB:SCI28.04c

Query Match 89.7%; Score 26; DB 2; Length 68;

Best Local Similarity 80.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
|:|:|
7 TRPPR 11

Db

RESULT 34

hypothetical protein lin1758 [imported] - Listeria innocua (strain Clp11262)

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C;Accession: AE1652

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloechl, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, I.; Jones, L.M.; Karst, U. Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vaquez-Boland, J.A.; Voss, H.; Wehlan, A.; Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AE1652

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-75 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC96989.1; PID:g16414245; GSPDB:GN00178

A;Experimental source: strain Clp11262

C;Genetics:

A;Gene: lin1758

Query Match 89.7%; Score 26; DB 2; Length 75;

Best Local Similarity 80.0%; Pred. No. 1.4e+02;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
|:|:|
38 TKPPK 42

Db

RESULT 35

conserved hypothetical protein YPO2363 [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001

C;Accession: AD0288

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AD0288

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-87 <KUR>

A;Cross-references: GB:AL590842; PIDN:CAC91168.1; PID:gl5980360; GSPDB:GN00175

C;Genetics:

A;Gene: YPO2363

Query Match 89.7%; Score 26; DB 2; Length 87;
 Best Local Similarity 80.0%; Pred. No. 1.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
 Db 83 TRPPR 87

RESULT 36
 T17560
 hypothetical protein a70L - Chlorella virus PBCV-1
 C;Species: Chlorella virus PBCV-1
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T17560
 R;Graves, M.V.; Van Etten, J.L.
 submitted to the EMBL Data Library, May 1999
 A;Reference number: Z18806
 A;Accession: T17560
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-88 <GRA>
 A;Cross-references: EMBL:U42580; NID:94028896; PIDN:AAC96438.1
 A;Experimental source: specific host Chlorella strain NC64A
 C;Genetics:
 A;Gene: a70L

Query Match 89.7%; Score 26; DB 2; Length 88;
 Best Local Similarity 80.0%; Pred. No. 1.6e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
 Db 51 TKPPK 55

RESULT 37
 F72587
 hypothetical protein APE1171 - Aeropyrum pernix (strain K1)
 C;Species: Aeropyrum pernix
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
 C;Accession: F72587
 R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Takahashi, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawa, H.; 83-101, 1999
 DNA Res. 6, 83-101, 1999
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
 A;Reference number: A72450; MUID:99310339; PMID:10382966
 A;Accession: F72587
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-100 <KAW>
 A;Cross-references: DDBJ:AF000061; NID:95104821; PIDN:BAAB0156.1; PID:1043942; PID:95104821
 A;Experimental source: strain K1
 C;Genetics:
 A;Gene: APE1171

Query Match 89.7%; Score 26; DB 2; Length 100;
 Best Local Similarity 80.0%; Pred. No. 1.8e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
 Db 70 TRPPR 74

RESULT 38
 JC2202
 secretin precursor - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 30-Sep-1993 #sequence_revision 20-Aug-1994 #text_change 16-Jul-1999
 C;Accession: JC2202; 834214
 R;Lian, M.S.; Kajiyama, W.; Donadel, G.; Lu, J.; Notkins, A.L.

Biochem. Biophys. Res. Commun. 200, 1066-1071, 1994
 A;Title: cDNA sequence and genomic organization of mouse secretin.
 A;Reference number: JC2202; MUID:94234995; PMID:8179583
 A;Accession: JC2202
 A;Molecule type: mRNA
 A;Residues: 1-133 <LAN>
 A;Cross-references: EMBL:X73580; NID:9313710; PIDN:CAA51982.1; PID:9313711
 C;Comment: This protein regulates the secretion of pancreatic juices and stimulates in C;Superfamily: glucagon
 C;Keywords: amidated carboxyl end; duplication; hormone; secretagogue
 F;1-27/Domain: signal sequence #status predicted <SIG>
 F;28-133/Product: prosecretin #status predicted <PRO>
 F;32-58/Product: secretin #status predicted <MAI>
 F;58/Modified site: amidated carboxyl end (Val) (amide in mature form from following 9

Query Match 89.7%; Score 26; DB 2; Length 133;
 Best Local Similarity 80.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
 Db 129 TRPPR 133

RESULT 39
 B49530
 vascular endothelial growth factor homolog A2R, 14.7K - Orf virus
 C;Species: Orf virus
 C;Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 08-Oct-1999
 C;Accession: B49530
 R;Lyttle, D.J.; Fraser, K.M.; Fleming, S.B.; Mercer, A.A.; Robinson, A.J.
 J. Virol. 68, 84-92, 1994
 A;Title: Homologs of vascular endothelial growth factor are encoded by the poxvirus or A;Reference number: A49530; MUID:94076465; PMID:8254780
 A;Contents: NZ2
 A;Accession: B49530
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-133 <LYT>
 A;Cross-references: GB:S67520; NID:9456897; PIDN:AAB29220.1; PID:9456899
 A;Note: sequence inconsistent with nucleotide translation
 A;Note: sequence extracted from NCBI backbone (NCBIN:141420, NCBIP:141425)

Query Match 89.7%; Score 26; DB 2; Length 133;
 Best Local Similarity 80.0%; Pred. No. 2.4e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
 Db 126 TRPPR 130

RESULT 40
 A29622
 ribosomal protein S12, mitochondrial - fruit fly (Drosophila melanogaster)
 N;Alternate names: technical knockout protein
 C;Species: Drosophila melanogaster
 C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 12-Jun-2003
 C;Accession: A29622
 R;Royden, C.S.; Pirrotta, V.; Jan, L.Y.
 Cell 51, 165-173, 1987
 A;Title: The tko locus, site of a behavioral mutation in D. melanogaster, codes for a A;Reference number: A29622; MUID:88027001; PMID:3117373
 A;Accession: A29622
 A;Molecule type: mRNA
 A;Residues: 1-140 <ROY>
 A;Cross-references: GB:M19494; NID:9158601; PIDN:AAA28935.1; PID:9158602
 C;Genetics:
 A;Gene: tko
 A;Cross-references: FlyBase:FBgn0003714
 C;Superfamily: ribosomal protein S12
 C;Keywords: mitochondrion; protein biosynthesis; ribosome
 F;117/Modified site: beta-methylthioaspartic acid (Asp) #status predicted

Query Match 89.7%; Score 26; DB 2; Length 140;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
|:|:|
Db 45 TRPPR 49

RESULT 41
T27059
hypothetical protein Y51A2A.6 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T27059
R/McMurray, A.
submitted to the EMBL Data Library, October 1998
A/Reference number: Z20304
A/Accession: T27059
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-140 <WIL>
A/Cross-references: EMBL:AL032635; PIDN:CRA21601.1; GSPDB:GN00023; CESP:Y51A2A.6
A/Experimental source: clone Y51A2A
C/Genetics:
A/Gene: CESP:Y51A2A.6
A/Map position: 5
A/Introns: 93/3; 129/1

Query Match 89.7%; Score 26; DB 2; Length 140;
Best Local Similarity 80.0%; Pred. No. 2.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
|:|:|
Db 68 TRPPR 72

RESULT 42
T49706
hypothetical protein B23L21.190 [imported] - Neurospora crassa
C/Species: Neurospora crassa
C/Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C/Accession: T49706
R/Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A/Reference number: Z25022
A/Accession: T49706
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-144 <SCH>
A/Cross-references: EMBL:AL356172; GSPDB:GN00116; NCSP:B23L21.190
A/Experimental source: BAC clone B23L21; strain OR74A
C/Genetics:
A/Gene: NCSP:B23L21.190
A/Map position: 6

Query Match 89.7%; Score 26; DB 2; Length 144;
Best Local Similarity 80.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
|:|:|
Db 130 TRPPR 134

RESULT 43
A87466
hypothetical protein C01749 [imported] - Caulobacter crescentus
C/Species: Caulobacter crescentus
C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C/Accession: A87466

R/Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Koln, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.I. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A/Title: Complete Genome Sequence of Caulobacter crescentus.
A/Reference number: A87249; MUID:21173698; PMID:11259647
A/Accession: A87466
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-148 <STO>
A/Cross-references: GB:AE005673; NID:gl3423171; PIDN:AAK23725.1; GSPDB:GN00148
C/Genetics:
A/Gene: C01749

Query Match 89.7%; Score 26; DB 2; Length 148;
Best Local Similarity 80.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
|:|:|
Db 82 TRPPR 86

RESULT 44
S75561
hypothetical protein sll0812 - Synecocystis sp. (strain PCC 6803)
C/Species: Synecocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C/Accession: S75561
R/Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasu DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocyst: s.
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S75561
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-153 <KAN>
A/Cross-references: EMBL:D90911; GB:AB001339; NID:gl653083; PIDN:BAA18122.1; PID:gl653;
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C/Superfamily: Synecocystis hypothetical protein sll0812

Query Match 89.7%; Score 26; DB 2; Length 153;
Best Local Similarity 80.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
|:|:|
Db 15 TRPPK 19

RESULT 45
A69267
transposase homolog - Archaeoglobus fulgidus
C/Species: Archaeoglobus fulgidus
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C/Accession: A69267
R/Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, F.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.I. Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997
A/Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S. Smith, H.O.; Woose, C.R.; Venter, J.C.
A/Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae
A/Reference number: A69250; MUID:98049343; PMID:9389475
A/Accession: A69267
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-154 <KLB>
A/Cross-references: GB:AE001097; GB:AE000782; NID:g2689420; PIDN:AA91094.1; PID:g26505

Query Match 89.7%; Score 26; DB 2; Length 154;
 Best Local Similarity 80.0%; Pred. No. 2.8e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 24 TRPPK 28

RESULT 46
 T49798
 conserved hypothetical protein [imported] - Neurospora crassa (fragment)
 N;Alternate names: protein B9J10.350
 C;Species: Neurospora crassa
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C;Accession: T49798
 R;Schulte, U.; Aign, V.; Hohisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
 submitted to the Protein Sequence Database, May 2000
 A;Reference number: 225022
 A;Accession: T49798
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-163 <SCH>
 A;Cross-references: EMBL:AL356324; GSPDB:GN00116; NCSP:B9J10.350
 A;Experimental source: SAC clone B9J10; strain OR74A
 C;Genetics:
 A;Gene: NCSP:B9J10.350
 A;Map position: 6

Query Match 89.7%; Score 26; DB 2; Length 163;
 Best Local Similarity 80.0%; Pred. No. 3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 45 TRPPR 49

RESULT 47
 AB1974
 hypothetical protein alr1341 [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C;Accession: AB1974
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi,
 Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
 baena sp. strain PCC 7120
 A;Reference number: AB1807; MUID:21595285; PMID:11759940
 A;Accession: AB1974
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-179 <KUR>
 A;Cross-references: GB:BA000019; PIDN:BA073298.1; PID:g17130688; GSPDB:GN00179
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: alr1341

Query Match 89.7%; Score 26; DB 2; Length 179;
 Best Local Similarity 80.0%; Pred. No. 3.3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 72 TRPPK 76

RESULT 48
 H83561
 hypothetical protein PA0671 [imported] - Pseudomonas aeruginosa (strain PAO1)
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C;Accession: H83561
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; E
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lin
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: H83561
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-183 <STO>
 A;Cross-references: GB:AE004502; GB:AE004091; NID:G9946547; PIDN:AA04060.1; GSPDB:GN01
 A;Experimental source: strain PAO1
 C;Genetics:
 A;Gene: PA0671

Query Match 89.7%; Score 26; DB 2; Length 183;
 Best Local Similarity 80.0%; Pred. No. 3.3e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 157 TRPPR 161

RESULT 49
 S45627
 acidic endoprotease precursor - Myxococcus xanthus (strain DK101)
 C;Species: Myxococcus xanthus
 A;Variety: strain DK101
 C;Date: 10-Dec-1994 #sequence_revision 26-Apr-1996 #text_change 08-Oct-1999
 C;Accession: S45627; S62857
 R;Lucas, N.; Mazaud-Aujard, C.; Bremaud, L.; Cenatiempo, Y.; Julien, R.
 Eur. J. Biochem. 222, 247-254, 1994
 A;Title: Protein purification, gene cloning and sequencing of an acidic endoprotease f
 A;Reference number: S45627; MUID:94291618; PMID:8020464
 A;Accession: S45627
 A;Molecule type: DNA
 A;Residues: 1-195 <LUC>
 A;Cross-references: EMBL:X75892; NID:G516391; PIDN:CAA53499.1; PID:G516392
 A;Experimental source: strain DK101
 A;Accession: S62857
 A;Molecule type: protein
 A;Residues: 65-101 <LUW>
 C;Genetics:
 A;Gene: Maep
 E;1-29/Domain: signal sequence #status predicted <SIG>
 E;30-64/Domain: propeptide #status predicted <PRO>
 F;65-195/Product: acidic endoprotease #status experimental <MAT>

Query Match 89.7%; Score 26; DB 2; Length 195;
 Best Local Similarity 80.0%; Pred. No. 3.5e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 148 TRPPR 152

RESULT 50
 FOA072
 major core protein VII precursor - human adenovirus 2
 C;Species: Mastadenovirus h2 (human adenovirus 2)
 A;Note: host Homo sapiens (man)
 C;Date: 18-Apr-1984 #sequence_revision 31-Dec-1993 #text_change 16-Jul-1999
 C;Accession: C03837; A03836
 R;Alestrom, P.; Akusjaervi, G.; Lager, M.; Yeh-kai, L.; Pettersson, U.
 J. Biol. Chem. 259, 13980-13985, 1984
 A;Title: Genes encoding the core proteins of adenovirus type 2.
 A;Reference number: A03837; MUID:85054835; PMID:6094534
 A;Accession: C03837
 A;Molecule type: DNA
 A;Residues: 1-198 <ALE>

A;Cross-references: GB:J01917; NID:G209811; PIDN:AAA92212.1; PID:G209828
R;Sung, M.T.; Cao, T.M.; Coleman, R.T.; Budellier, K.A.
Proc. Natl. Acad. Sci. U.S.A. 80, 2902-2906, 1983
A;Title: Gene and protein sequences of adenovirus protein VII, a hybrid basic chromosoma
A;Reference number: A03836; MUID:83221511; PMID:6574459
A;Accession: A03836
A;Molecule type: DNA
A;Residues: 1-111/113-198 <SUN>
A;Cross-references: GB:J01917
C;Genetics:
A;Map position: 43-45
C;Superfamily: adenovirus major core protein VII
C;Keywords: core protein, late protein
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-198/Product: major core protein VII #status predicted <MCP>

Query Match 89.7%; Score 26; DB 1; Length 198;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPR 5
|:|
Db 193 TRPR 197

Search completed: March 3, 2004, 12:19:05
Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 12:16:13 ; Search time 23 seconds
(without alignments)
11.223 Million cell updates/sec

Title: US-09-871-974-2

Perfect score: 29

Sequence: 1 TKPPR 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCITUS COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	29	100.0	5	1	US-08-171-737-2
2	29	100.0	5	1	US-08-202-178-2
3	29	100.0	5	1	US-08-299-636-37
4	29	100.0	5	1	US-08-279-155-36
5	29	100.0	5	1	US-08-713-484-2
6	29	100.0	5	1	US-08-703-988A-36
7	29	100.0	5	1	US-08-532-294-1
8	29	100.0	5	1	US-08-454-859-1
9	29	100.0	5	2	US-08-955-263-2
10	29	100.0	5	2	US-08-612-842-36
11	29	100.0	6	1	US-08-171-737-18
12	29	100.0	6	1	US-08-299-636-38
13	29	100.0	6	1	US-08-279-155-37
14	29	100.0	6	1	US-08-703-988A-37
15	29	100.0	6	2	US-08-612-842-37
16	29	100.0	8	1	US-08-202-178-4
17	29	100.0	8	1	US-08-713-484-4
18	29	100.0	8	1	US-08-532-294-2
19	29	100.0	8	2	US-08-955-263-4
20	29	100.0	8	2	US-08-612-842-39
21	29	100.0	9	1	US-08-171-737-9
22	29	100.0	9	1	US-08-279-155-1
23	29	100.0	9	1	US-08-279-155-38
24	29	100.0	9	1	US-08-703-988A-1
25	29	100.0	9	1	US-08-703-988A-38
26	29	100.0	9	1	US-08-532-294-3
27	29	100.0	9	1	US-08-454-859-2

28	29	100.0	9	2	US-08-613-842-1	Sequence 1, Appl
29	29	100.0	9	2	US-08-613-842-38	Sequence 38, Appl
30	29	100.0	9	4	US-08-997-802-10	Sequence 10, Appl
31	29	100.0	9	4	US-08-997-802-11	Sequence 11, Appl
32	29	100.0	10	1	US-08-202-178-5	Sequence 5, Appl
33	29	100.0	10	1	US-08-713-484-5	Sequence 5, Appl
34	29	100.0	10	1	US-08-955-263-5	Sequence 5, Appl
35	29	100.0	11	1	US-08-202-178-7	Sequence 7, Appl
36	29	100.0	11	1	US-08-713-484-7	Sequence 7, Appl
37	29	100.0	11	2	US-08-955-263-7	Sequence 7, Appl
38	29	100.0	11	4	US-09-387-715-51	Sequence 51, Appl
39	29	100.0	11	4	US-09-387-715-52	Sequence 52, Appl
40	29	100.0	12	1	US-08-299-636-39	Sequence 39, Appl
41	29	100.0	60	4	US-09-381-546-13	Sequence 13, Appl
42	29	100.0	98	4	US-09-252-991A-25421	Sequence 25421, A
43	29	100.0	136	4	US-09-252-991A-21156	Sequence 21156, A
44	29	100.0	420	4	US-09-489-039A-8820	Sequence 8820, Ap
45	29	100.0	576	4	US-09-367-206-1	Sequence 1, Appl
46	29	100.0	576	4	US-09-367-206-21	Sequence 21, Appl
47	29	100.0	576	4	US-09-367-206-22	Sequence 22, Appl
48	29	100.0	576	4	US-09-367-206-23	Sequence 23, Appl
49	29	100.0	892	3	US-08-857-076-42	Sequence 42, Appl
50	29	100.0	1235	1	US-08-118-101A-2	Sequence 2, Appl
51	26	89.7	14	4	US-09-608-285A-44	Sequence 44, Appl
52	26	89.7	14	4	US-09-557-800C-44	Sequence 44, Appl
53	26	89.7	29	4	US-09-287-849-30	Sequence 30, Appl
54	26	89.7	44	4	US-09-645-470-14	Sequence 14, Appl
55	26	89.7	67	4	US-09-107-532A-3781	Sequence 3781, Ap
56	26	89.7	128	4	US-09-252-991A-28839	Sequence 28839, A
57	26	89.7	132	4	US-09-125-642C-15	Sequence 15, Appl
58	26	89.7	132	4	US-09-431-888-11	Sequence 11, Appl
59	26	89.7	133	4	US-09-732-210-1700	Sequence 2, Appl
60	26	89.7	140	4	US-09-252-991A-18713	Sequence 18713, A
61	26	89.7	147	4	US-09-252-991A-18672	Sequence 18672, A
62	26	89.7	178	4	US-09-252-991A-36873	Sequence 36873, A
63	26	89.7	211	4	US-09-252-991A-30029	Sequence 30029, A
64	26	89.7	211	4	US-09-252-991A-16739	Sequence 16739, A
65	26	89.7	220	4	US-09-252-991A-28621	Sequence 28621, A
66	26	89.7	234	4	US-09-252-991A-32366	Sequence 32366, A
67	26	89.7	245	4	US-09-252-991A-20329	Sequence 20329, A
68	26	89.7	253	4	US-08-565-386-9	Sequence 9, Appl
69	26	89.7	263	1	US-09-252-991A-22289	Sequence 22289, A
70	26	89.7	264	4	US-09-252-991A-19978	Sequence 19978, A
71	26	89.7	282	4	US-09-564-951B-4	Sequence 4, Appl
72	26	89.7	290	4	US-09-355-040-13	Sequence 13, Appl
73	26	89.7	307	4	US-09-355-040-14	Sequence 14, Appl
74	26	89.7	307	4	US-09-050-739-94	Sequence 94, Appl
75	26	89.7	308	4	US-09-252-991A-17744	Sequence 17744, A
76	26	89.7	309	4	US-09-540-236-2724	Sequence 2724, Ap
77	26	89.7	313	3	US-08-612-973-42	Sequence 42, Appl
78	26	89.7	314	3	US-08-927-597-42	Sequence 42, Appl
79	26	89.7	316	4	US-09-252-991A-30881	Sequence 30881, A
80	26	89.7	316	4	US-08-612-973-44	Sequence 44, Appl
81	26	89.7	319	3	US-08-927-597-44	Sequence 44, Appl
82	26	89.7	319	3	US-09-252-991A-34925	Sequence 34925, A
83	26	89.7	321	4	US-09-543-681A-5632	Sequence 5632, Ap
84	26	89.7	337	3	US-08-612-973-38	Sequence 38, Appl
85	26	89.7	338	3	US-08-927-597-38	Sequence 38, Appl
86	26	89.7	343	3	US-08-612-973-40	Sequence 40, Appl
87	26	89.7	343	3	US-08-927-597-40	Sequence 40, Appl
88	26	89.7	343	4	US-09-252-991A-30578	Sequence 30578, A
89	26	89.7	350	4	US-09-818-780-95	Sequence 95, Appl
90	26	89.7	350	4	US-09-269-137-8	Sequence 8, Appl
91	26	89.7	359	4	US-09-355-040-30	Sequence 30, Appl
92	26	89.7	369	4	US-09-540-236-2085	Sequence 2085, Ap
93	26	89.7	375	4	US-09-252-991A-18786	Sequence 18786, A
94	26	89.7	384	4	US-09-107-532A-3712	Sequence 3712, Ap
95	26	89.7	393	3	US-08-928-442-1	Sequence 1, Appl
96	26	89.7	395	3	US-08-570-157-5	Sequence 5, Appl
97	26	89.7	428	3	US-08-029-170-31	Sequence 31, Appl
98	26	89.7	428	3	US-09-076-510-5	Sequence 5, Appl
99	26	89.7	428	4	US-09-004-349-5	Sequence 5, Appl
100	26	89.7	428	4		

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPR 5
| | | | |
Db 1 TKPR 5

RESULT 3
US-08-299-636-37
; Sequence 37, Application US/08299636
; Patent No. 5659041
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: KIRBY, Robert A.
; APPLICANT: DUNN-DUFAULT, Robert
; TITLE OF INVENTION: HYDRAZINO-TYPE RADIONUCLIDE CHELATORS
; TITLE OF INVENTION: HAVING AN N3S CONFIGURATION
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,636
; FILING DATE: 02-SEP-1994
; CLASSIFICATION: 534
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/092,911
; FILING DATE: 18-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/262/ALLE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-299-636-37

Query Match 100.0%; Score 29; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPR 5
| | | | |
Db 1 TKPR 5

RESULT 4
US-08-279-155-36
; Sequence 36, Application US/08279155
; Patent No. 5662885
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: GOODBODY, Anne
; TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIKO, MARWELSTEIN, MURRAY & ORAM LLP

STREET: 655 Fifteenth Street, N. W., Suite 330 - G
STREET: Street Lobby
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/279,155
FILING DATE: 22-JUL-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MURRAY, Robert B.
REGISTRATION NUMBER: 22,980
REFERENCE/DOCKET NUMBER: P8074-4005
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 5 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-279-155-36

Query Match 100.0%; Score 29; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPR 5
| | | | |
Db 1 TKPR 5

RESULT 5
US-08-713-484-2
; Sequence 2, Application US/08713484
; Patent No. 5679642
; GENERAL INFORMATION:
; APPLICANT: Goodbody, Anne
; APPLICANT: POLLAK, Alfred
; TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713,484
FILING DATE: 13-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/202,178
FILING DATE: 25-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION Number: 29,768
REFERENCE/DOCKET NUMBER: 16777/290/ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399

AC

; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,859
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,107
; REFERENCE/DOCKET NUMBER: 8074-5007
; TELEPHONE: (202) 638-5000
; TELEFAX: (202) 638-4810
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-454-859-1

Query Match 100.0%; Score 29; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 TKPPR 5
Db 1 TKPPR 5

RESULT 9
US-08-955-263-2
; Sequence 2, Application US/08955263
; Patent No. 5866544
; GENERAL INFORMATION:
; APPLICANT: Goodbody, Anne
; APPLICANT: Pollak, Alfred
; TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,263
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/713,484
; FILING DATE: 13-SEP-1996
; APPLICATION NUMBER: US 08/202,178
; FILING DATE: 25-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/290/ALLE
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-955-263-2

Query Match 100.0%; Score 29; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 TKPPR 5
Db 1 TKPPR 5

RESULT 10
US-08-612-842-36
; Sequence 36, Application US/08612842
; Patent No. 5976495
; GENERAL INFORMATION:
; APPLICANT: POLLAK, ALFRED
; APPLICANT: GOODBODY, ANNE
; TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAI DO, MARCELSTEIN, MURRAY & ORAM
; STREET: 655 15TH STREET, NW, G STREET LOBBY, SUITE
; STREET: 330
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,842
; FILING DATE: 20-MAR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: BERMAN, RICHARD J.
; REGISTRATION NUMBER: 39107
; REFERENCE/DOCKET NUMBER: 8012-6002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 638-5000
; TELEFAX: 202 638-4810
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-612-842-36

Query Match 100.0%; Score 29; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 3e+05; Indels 0;
Matches 5; Conservative 0; Mismatches 0; Gaps 0;

Qy 1 TKPPR 5
Db 1 TKPPR 5

RESULT 11
US-08-171-737-18
; Sequence 18, Application US/08171737
; Patent No. 5480970
; GENERAL INFORMATION:

APPLICANT: Pollak, Alfred
APPLICANT: Goodbody, Anne
TITLE OF INVENTION: METAL CHELATORS
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/171,737
FILING DATE: 22-DEC-1993
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/253/ALLE
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-171-737-18

Query Match 100.0%; Score 29; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 2 TKPPR 6

RESULT 12
US-08-299-636-38
; Sequence 38, Application US/08299636
; Patent No. 565904L
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: KIRBY, Robert A.
; APPLICANT: DUNN-DUFAULT, Robert
; TITLE OF INVENTION: HYDRAZINO-TYPE RADIONUCLIDE CHELATORS
; TITLE OF INVENTION: HAVING AN N3S CONFIGURATION
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,636
; FILING DATE: 02-SEP-1994
; CLASSIFICATION: 534
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/092,911

FILING DATE: 18-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/262/ALLE
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 6 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-299-636-38

Query Match 100.0%; Score 29; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 2 TKPPR 6

RESULT 13
US-08-279-155-37
; Sequence 37, Application US/08279155
; Patent No. 5662885
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: GOODBODY, Anne
; TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NAKAIDO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W., Suite 330 - G
; STREET: Street Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/279,155
; FILING DATE: 22-JUL-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MURRAY, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P8074-4005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6
; OTHER INFORMATION: /note= "Arg at position 6 has an OH
; OTHER INFORMATION: group."
; US-08-279-155-37

Query Match 100.0%; Score 29; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 2 TKPPR 6

RESULT 14

US-08-703-988A-37
; Sequence 37, Application US/08703988A
; Patent No. 5780006
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: GOODBODY, Anne
; TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE
; TITLE OF INVENTION: CHELATORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAI DO, MARMELESTEIN, MURRAY & ORAM
; ADDRESSEE: LLP
; STREET: 655 Fifteenth Street, N. W., Suite 330
; STREET:
; STREET: G Street Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/703,988A
; FILING DATE: 28-AUG-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/279,155
; FILING DATE: 22-JUL-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MURRAY, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P8074-6011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202/638-5000
; TELEFAX: 202/638-4810
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6
; OTHER INFORMATION: /note= "Arg at position 6 has
; OTHER INFORMATION: an OH group."
US-08-703-988A-37

Query Match 100.0%; Score 29; DB 1; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 2 TKPPR 6

RESULT 15

US-08-612-842-37
; Sequence 37, Application US/08612842
; Patent No. 5976495
; GENERAL INFORMATION:
; APPLICANT: POLLAK, ALFRED
; APPLICANT: GOODBODY, ANNE
; TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAI DO, MARMELESTEIN, MURRAY & ORAM
; STREET: 655 15TH STREET, NW, G STREET LOBBY, SUITE
; STREET: 330
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,842
; FILING DATE: 20-MAR-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: BERMAN, RICHARD J.
; REGISTRATION NUMBER: 39107
; REFERENCE/DOCKET NUMBER: 8012-6002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 638-5000
; TELEFAX: 202 638-4810
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6
; OTHER INFORMATION: /note= "Position 6 has an -OH
; OTHER INFORMATION: substituent."
US-08-612-842-37

Query Match 100.0%; Score 29; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 2 TKPPR 6

RESULT 16

US-08-202-178-4
; Sequence 4, Application US/08202178
; Patent No. 5569745
; GENERAL INFORMATION:
; APPLICANT: Goodbody, Anne
; APPLICANT: Pollak, Alfred
; TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,178
; FILING DATE: 25-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/258/ALLE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Ser substituted with
; OTHER INFORMATION: picolinic acid (Pic)."
```

US-08-202-178-4

```
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /note= "Cys substituted with
; OTHER INFORMATION: acetamidomethyl (AcM)."
```

Query Match 100.0%; Score 29; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 4 TKPPR 8

RESULT 17

US-08-713-484-4

```
; Sequence 4, Application US/08713484
; Patent No. 5679642
; GENERAL INFORMATION:
; APPLICANT: Goodbody, Anne
; APPLICANT: Pollak, Alfred
; TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,484
; FILING DATE: 13-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/202,178
; FILING DATE: 25-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/290/ALLE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
```

```
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Ser substituted with
; OTHER INFORMATION: picolinic acid (Pic)."
```

US-08-713-484-4

```
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /note= "Cys substituted with
; OTHER INFORMATION: acetamidomethyl (AcM)."
```

Query Match 100.0%; Score 29; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 4 TKPPR 8

RESULT 18

US-08-592-294-2

```
; Sequence 2, Application US/08592294
; Patent No. 5789555
; GENERAL INFORMATION:
; APPLICANT: POLLAK, ALFRED
; APPLICANT: DUNN-DUFAULT, ROBERT
; TITLE OF INVENTION: IMMOBILIZED LABELLING METHOD
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKALDO, MARKELSTEIN, MURRAY & ORAM LLP
; STREET: 655 FIFTEENTH ST., N.W., SUITE 330-G STREET
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/592,294
; FILING DATE:
; CLASSIFICATION: 534
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/CA94/00637
; FILING DATE: 16-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/152,680
; FILING DATE: 16-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: MURRAY, ROBERT B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P8012-5002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-638-5000
; TELEFAX: 202-638-4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
```

TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1
OTHER INFORMATION: /note= "Picolinic acid is attached to Ser of position 1."
US-08-592-294-2

Query Match 100.0%; Score 29; DB 1; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 4 TKPPR 8

RESULT 19
US-08-955-263-4
Sequence 4, Application US/08955263
Patent No. 5866544
GENERAL INFORMATION:
APPLICANT: Goodbody, Anne
APPLICANT: Pollak, Alfred
TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,263
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/713,484
FILING DATE: 13-SEP-1996
APPLICATION NUMBER: US 08/202,178
FILING DATE: 25-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/290/ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Ser substituted with picolinic acid (Pic)."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2
OTHER INFORMATION: /note= "Cys substituted with acetamidomethyl (Acm)."
US-08-955-263-4
Query Match 100.0%; Score 29; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TKPPR 5
Db 4 TKPPR 8

RESULT 20
US-08-612-842-39
Sequence 39, Application US/08612842
Patent No. 5976495
GENERAL INFORMATION:
APPLICANT: POLLAK, ALFRED
APPLICANT: GOODBODY, ANNE
TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKAIIDO, MARCELSTEIN, MURRAY & ORAM
STREET: 655 15TH STREET, NW, G STREET LOBBY, SUITE
STREET: 330
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612,842
FILING DATE: 20-MAR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BERMAN, RICHARD J
REGISTRATION NUMBER: 39107
REFERENCE/DOCKET NUMBER: 8012-6002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 638-5000
TELEFAX: 202 638-4810
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Position 1 has a sarcosine substituent."
US-08-612-842-39

Query Match 100.0%; Score 29; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 4 TKPPR 8

RESULT 21
US-08-171-737-9
Sequence 9, Application US/08171737
Patent No. 5480970
GENERAL INFORMATION:
APPLICANT: Pollak, Alfred
APPLICANT: Goodbody, Anne
TITLE OF INVENTION: METAL CHELATORS
NUMBER OF SEQUENCES: 19

;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: Suite 500, 3000 K Street N.W.
;; CITY: Washington, D.C.
;; COUNTRY: USA
;; ZIP: 20007-5109
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent In Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/171,737
;; FILING DATE: 22-DEC-1993
;; CLASSIFICATION: 514
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Bent, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 16777/253/ALLE
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 672-5300
;; TELEFAX: (202) 672-5399
;; TELEX: 904136
;; INFORMATION FOR SEQ ID NO: 9:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: peptide
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 1
;; OTHER INFORMATION: /note= "X at position 1 is
;; OTHER INFORMATION: picolinic acid (Pic)."
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 3
;; OTHER INFORMATION: /note= "Cys at position 3 is
;; OTHER INFORMATION: substituted with acetamidomethyl (Acm)."
US-08-171-737-9
Query Match 100.0%; Score 29; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
DB 5 TKPPR 9
RESULT 22
US-08-279-155-1
; Sequence 1, Application US/08279155
; Patent No. 5662885
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: GOODBODY, Anne
; TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W., Suite 330 - G
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/279,155
;; FILING DATE: 22-JUL-1994
;; CLASSIFICATION: 424
;; ATTORNEY/AGENT INFORMATION:
;; NAME: MURRAY, Robert B.
;; REGISTRATION NUMBER: 22,980
;; REFERENCE/DOCKET NUMBER: P8074-4005
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 202/638-5000
;; TELEFAX: 202/638-4810
;; INFORMATION FOR SEQ ID NO: 1:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 9 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 1
;; OTHER INFORMATION: /note= "Gly at position 1 has an
;; OTHER INFORMATION: N,N-dimethyl group."
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 3
;; OTHER INFORMATION: /note= "Cys at position 3 has an
;; OTHER INFORMATION: Acm group."
;; FEATURE:
;; NAME/KEY: Modified-site
;; LOCATION: 9
;; OTHER INFORMATION: /note= "Arg at position 10 is
;; OTHER INFORMATION: unsubstituted or has an OH group."
US-08-279-155-1
Query Match 100.0%; Score 29; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
DB 5 TKPPR 9
RESULT 23
US-08-279-155-38
; Sequence 38, Application US/08279155
; Patent No. 5663885
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: GOODBODY, Anne
; TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth Street, N. W., Suite 330 - G
; CITY: Street Lobby
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/279,155
; FILING DATE: 22-JUL-1994
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: MURRAY, Robert B.
; REGISTRATION NUMBER: 22,980
; REFERENCE/DOCKET NUMBER: P8074-4005

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Gly at position 1 has an N,N-dimethyl group."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /note= "Cys at position 3 has an Acn group."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note= "Arg at position 9 is unsubstituted or has an OH group."
US-08-279-155-38

Query Match 100.0%; Score 29; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 5 TKPPR 9

RESULT 24
US-08-703-988A-1
Sequence 1, Application US/08703988A
Patent No. 5780006
GENERAL INFORMATION:
APPLICANT: POLLAK, Alfred
APPLICANT: GOODBODY, Anne
TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE
TITLE OF INVENTION: CHELATORS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM
ADDRESSEE: LLP
STREET: 655 Fifteenth Street, N. W., Suite 330
STREET: G Street Lobby
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/703,988A
FILING DATE: 28-AUG-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/279,155
FILING DATE: 22-JUL-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MURRAY, Robert B.
REGISTRATION NUMBER: 22,980
REFERENCE/DOCKET NUMBER: P8074-6011

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Gly at position 1 has an N,N-dimethyl group."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /note= "Cys at position 3 has an Acn group."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note= "Arg at position 9 is unsubstituted or has an OH group."
US-08-703-988A-1

Query Match 100.0%; Score 29; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 5 TKPPR 9

RESULT 25
US-08-703-988A-38
Sequence 38, Application US/08703988A
Patent No. 5780006
GENERAL INFORMATION:
APPLICANT: POLLAK, Alfred
APPLICANT: GOODBODY, Anne
TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE
TITLE OF INVENTION: CHELATORS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM
ADDRESSEE: LLP
STREET: 655 Fifteenth Street, N. W., Suite 330
STREET: G Street Lobby
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/703,988A
FILING DATE: 28-AUG-1996
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/279,155
FILING DATE: 22-JUL-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: MURRAY, Robert B.
REGISTRATION NUMBER: 22,980
REFERENCE/DOCKET NUMBER: P8074-6011

TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000
TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "gly at position 1 has
OTHER INFORMATION: an N,N-dimethyl group."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 3
OTHER INFORMATION: /note= "Cys at position 3 has
OTHER INFORMATION: an AcM group."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note= "Arg at position 9 is
OTHER INFORMATION: unsubstituted or has an OH group."

US-08-703-988A-38

Query Match 100.0%; Score 29; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 5 TKPPR 9

RESULT 26

US-08-592-294-3
Sequence 3, Application US/08592294
Patent No. 5789555
GENERAL INFORMATION:
APPLICANT: POLLAK, ALFRED
APPLICANT: DUNN-DUFAULT, ROBERT
TITLE OF INVENTION: IMMOBILIZED LABELLING METHOD
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
STREET: 655 FIFTEENTH ST., N.W., SUITE 330-G STREET
STREET: LOBBY
CITY: WASHINGTON
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,294
FILING DATE:
CLASSIFICATION: 534
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/CA94/00637
FILING DATE: 16-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/152,680
FILING DATE: 16-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: MURRAY, ROBERT B
REGISTRATION NUMBER: 22,980
REFERENCE/DOCKET NUMBER: P8012-5002
TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-638-5000
TELEFAX: 202-638-4810
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1
OTHER INFORMATION: /note= "N,N'-dimethyl is attached to Gly
OTHER INFORMATION: of position 1."
US-08-592-294-3

Query Match 100.0%; Score 29; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 5 TKPPR 9

RESULT 27

US-08-454-859-2
Sequence 2, Application US/08454859
Patent No. 580458
GENERAL INFORMATION:
APPLICANT: POLLAK, Alfred
TITLE OF INVENTION: SEQUESTERED IMAGING AGENTS
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM
STREET: 655 FIFTEENTH STREET, NW, SUITE 330, G.
STREET: Street Lobby
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,859
FILING DATE: 31-MAY-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Bertan, Richard J.
REGISTRATION NUMBER: 39,107
REFERENCE/DOCKET NUMBER: 8074-5007
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638-5000
TELEFAX: (202) 638-4810
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Position 1 has either a
OTHER INFORMATION: dimethyl substituent or an N,N'-dimethyl substituent."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note= "Position 9 has an -OH

OTHER INFORMATION: substituent."

-454-859-2

ry Match 100.0%; Score 29; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05; 0; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TKPR 5
|||||
5 TKPR 9

1T 28

8-612-842-1
Sequence 1, Application US/08612842
Patent No. 5976495
GENERAL INFORMATION:
APPLICANT: POLLAK, ALFRED
APPLICANT: GOODBODY, ANNE
TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKALDO, MARCELSTEIN, MURRAY & ORAM
STREET: 655 15TH STREET, NW, G STREET LOBBY, SUITE
STREET: 330
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612.842
FILING DATE: 20-MAR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BERNAN, RICHARD J
REGISTRATION NUMBER: 39107
REFERENCE/DOCKET NUMBER: 8012-6002
TELEPHONE: 202 638-4810
TELEFAX: 202 638-5000
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Position 1 has either an N,N-dibenzyl, an N,N-diethyl, or a sarcosine substituent."

1 TKPR 5
|||||
5 TKPR 9

Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local similarity 100.0%; Pred. No. 3e+05; 0; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TKPR 5
|||||
5 TKPR 9

RESULT 29
US-08-612-842-38
Sequence 38, Application US/08612842
Patent No. 5976495
GENERAL INFORMATION:
APPLICANT: POLLAK, ALFRED
APPLICANT: GOODBODY, ANNE
TITLE OF INVENTION: PEPTIDE DERIVED RADIONUCLIDE CHELATORS
NUMBER OF SEQUENCES: 39
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKALDO, MARCELSTEIN, MURRAY & ORAM
STREET: 655 15TH STREET, NW, G STREET LOBBY, SUITE
STREET: 330
CITY: WASHINGTON
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/612.842
FILING DATE: 20-MAR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: BERNAN, RICHARD J
REGISTRATION NUMBER: 39107
REFERENCE/DOCKET NUMBER: 8012-6002
TELEPHONE: 202 638-5000
TELEFAX: 202 638-4810
INFORMATION FOR SEQ ID NO: 38:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Position 1 has an N,N-dimethyl substituent."

1 TKPR 5
|||||
5 TKPR 9

Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05; 0; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 30
US-08-997-802-10
Sequence 10, Application US/08997802
Patent No. 6334996
GENERAL INFORMATION:
APPLICANT: WONG, Ernest
TITLE OF INVENTION: CHELATORS THAT PREDOMINATELY FORM A SINGLE
TITLE OF INVENTION: STEREOISOMERIC SPECIES UPON COORDINATION TO A METAL
FILE REFERENCE: 8298-7019
CURRENT APPLICATION NUMBER: US/08/997.802

CURRENT FILING DATE: 1997-12-24
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10

LENGTH: 9
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (1)_RES
OTHER INFORMATION: dimethylglycine
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (2)_
OTHER INFORMATION: L-t-butylglycine
FEATURE:
NAME/KEY: MOD_RES
LOCATION: (3)_
OTHER INFORMATION: L-cysteine with an acetoamidomethyl protecting
OTHER INFORMATION: group attached via the Sulfur atom.
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
OTHER INFORMATION: Peptide
S-08-997-802-10

Query Match 100.0%; Score 29; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 TKPR 5
|||
Db 5 TKPR 9

RESULT 31
JS-08-997-802-11
; Sequence 11, Application US/08997802
; Patent No. 6334996
; GENERAL INFORMATION:
; APPLICANT: WONG, Ernest
; TITLE OF INVENTION: CHELATORS THAT PREDOMINATELY FORM A SINGLE
; TITLE OF INVENTION: STEREOISOMERIC SPECIES UPON COORDINATION TO A METAL
; TITLE OF INVENTION: CENTER
; FILE REFERENCE: 8298-7019
; CURRENT APPLICATION NUMBER: US/08/997,802
; CURRENT FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)_
; OTHER INFORMATION: Dimethylglycine.
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (2)_
; OTHER INFORMATION: L-t-butyl glycine
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (3)_
; OTHER INFORMATION: L-cysteine.
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-08-997-802-11

Query Match 100.0%; Score 29; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPR 5
|||
Db 5 TKPR 9

RESULT 32
US-08-202-178-5
; Sequence 5, Application US/08202178
; Patent No. 5569745
; GENERAL INFORMATION:
; APPLICANT: Goodbody, Anne
; APPLICANT: Pollak, Alfred
; TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,178
; FILING DATE: 25-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/258/ALLE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /note= "Cys substituted with
; OTHER INFORMATION: acetamidomethyl (AcM)."
US-08-202-178-5

Query Match 100.0%; Score 29; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPR 5
|||
Db 6 TKPR 10

RESULT 33
US-08-713-484-5
; Sequence 5, Application US/08713484
; Patent No. 5679642
; GENERAL INFORMATION:
; APPLICANT: Goodbody, Anne
; APPLICANT: Pollak, Alfred
; TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES

NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/713,484
FILING DATE: 13-SEP-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/202,178
FILING DATE: 25-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/290/ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Ser substituted with
OTHER INFORMATION: picolinic acid (Pic)."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2
OTHER INFORMATION: /note= "Cys substituted with
OTHER INFORMATION: acetamidomethyl (Acm)."
US-08-713-484-5

Query Match 100.0%; Score 29; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TKPPR 5
Db 6 TKPPR 10

RESULT 34
US-08-955-263-5
Sequence 5, Application US/08955263
Patent No. 5866544
GENERAL INFORMATION:
APPLICANT: Goodbody, Anne
APPLICANT: Pollak, Alfred
TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/955,263
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/713,484
FILING DATE: 13-SEP-1996
APPLICATION NUMBER: US 08/202,178
FILING DATE: 25-FEB-1994
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/290/ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "Ser substituted with
OTHER INFORMATION: picolinic acid (Pic)."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 2
OTHER INFORMATION: /note= "Cys substituted with
OTHER INFORMATION: acetamidomethyl (Acm)."
US-08-955-263-5

Query Match 100.0%; Score 29; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TKPPR 5
Db 6 TKPPR 10

RESULT 35
US-08-202-178-7
Sequence 7, Application US/08202178
Patent No. 5569745
GENERAL INFORMATION:
APPLICANT: Goodbody, Anne
APPLICANT: Pollak, Alfred
TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: Suite 500, 3000 K Street, N.W.
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/202,178
FILING DATE: 25-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bent, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16777/258/ALLE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300

```
;
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Ser is substituted with
; OTHER INFORMATION: benzoylmercaptoacetic acid (Bz-MA)."
;
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /note= "Cys is substituted with
; OTHER INFORMATION: acetamidomethyl (Acm)."
;
US-08-202-178-7

Query Match 100.0%; Score 29; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPR 5
Db 7 TKPR 11

RESULT 36
US-08-713-484-7
; Sequence 7, Application US/08713484
; Patent No. 5679642
; GENERAL INFORMATION:
; APPLICANT: Goodbody, Anne
; APPLICANT: Pollak, Alfred
; TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,484
; FILING DATE: 13-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/202,178
; FILING DATE: 25-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/290/ALLE
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Ser is substituted with
; OTHER INFORMATION: benzoylmercaptoacetic acid (Bz-MA)."
;
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /note= "Cys is substituted with
; OTHER INFORMATION: acetamidomethyl (Acm)."
;
US-08-955-263-7

Query Match 100.0%; Score 29; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPR 5
Db 7 TKPR 11

RESULT 37
US-08-955-263-7
; Sequence 7, Application US/08955263
; Patent No. 5866544
; GENERAL INFORMATION:
; APPLICANT: Goodbody, Anne
; APPLICANT: Pollak, Alfred
; TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,263
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/713,484
; FILING DATE: 13-SEP-1996
; APPLICATION NUMBER: US 08/202,178
; FILING DATE: 25-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/290/ALLE
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Ser is substituted with
; OTHER INFORMATION: benzoylmercaptoacetic acid (Bz-MA)."
;
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /note= "Cys is substituted with
; OTHER INFORMATION: acetamidomethyl (Acm)."
;
US-08-955-263-7

Query Match 100.0%; Score 29; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
```

```
;
; OTHER INFORMATION: benzoylmercaptoacetic acid (Bz-MA)."
;
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /note= "Cys is substituted with
; OTHER INFORMATION: acetamidomethyl (Acm)."
;
US-08-713-484-7

Query Match 100.0%; Score 29; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPR 5
Db 7 TKPR 11

RESULT 37
US-08-955-263-7
; Sequence 7, Application US/08955263
; Patent No. 5866544
; GENERAL INFORMATION:
; APPLICANT: Goodbody, Anne
; APPLICANT: Pollak, Alfred
; TITLE OF INVENTION: PEPTIDE-CHELATOR CONJUGATES
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/955,263
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/713,484
; FILING DATE: 13-SEP-1996
; APPLICATION NUMBER: US 08/202,178
; FILING DATE: 25-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bent, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/290/ALLE
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /note= "Ser is substituted with
; OTHER INFORMATION: benzoylmercaptoacetic acid (Bz-MA)."
;
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /note= "Cys is substituted with
; OTHER INFORMATION: acetamidomethyl (Acm)."
;
US-08-955-263-7

Query Match 100.0%; Score 29; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
| | | | |
Db 7 TKPPR 11

RESULT 38

US-09-387-715-51
; Sequence 51, Application US/09387715
; Patent No. 6551574
; GENERAL INFORMATION:
; APPLICANT: Sharma, Shubh
; TITLE OF INVENTION: Tuftsin Metallopeptides Analogs and Uses Thereof
; FILE REFERENCE: 1173/1D794US1
; CURRENT APPLICATION NUMBER: US/09/387,715
; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: PCT/US99/05693
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 51
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: MOD_RES
; LOCATION: 6
; OTHER INFORMATION: Ahe
; NAME/KEY: modified residue
; LOCATION: 8,10
; OTHER INFORMATION: D-amino acid
US-09-387-715-51

Query Match 100.0%; Score 29; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
| | | | |
Db 1 TKPPR 5

RESULT 39

US-09-387-715-52
; Sequence 52, Application US/09387715
; Patent No. 6551574
; GENERAL INFORMATION:
; APPLICANT: Sharma, Shubh
; TITLE OF INVENTION: Tuftsin Metallopeptides Analogs and Uses Thereof
; FILE REFERENCE: 1173/1D794US1
; CURRENT APPLICATION NUMBER: US/09/387,715
; CURRENT FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: PCT/US99/05693
; PRIOR FILING DATE: 1999-03-18
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 52
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: peptide
; NAME/KEY: MOD_RES
; LOCATION: 6
; OTHER INFORMATION: Ahe
; NAME/KEY: modified residue
; LOCATION: 2,4
; OTHER INFORMATION: D-amino acid
US-09-387-715-52

Query Match 100.0%; Score 29; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
| | | | |
Db 7 TKPPR 11

RESULT 40

US-08-299-636-39
; Sequence 39, Application US/08299636
; Patent No. 5659041
; GENERAL INFORMATION:
; APPLICANT: POLLAK, Alfred
; APPLICANT: KIRBY, Robert A.
; APPLICANT: DUNN-DUFAULT, Robert
; TITLE OF INVENTION: HYDRAZINO-TYPE RADIONUCLIDE CHELATORS
; TITLE OF INVENTION: HAVING AN NIS CONFIGURATION
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D C
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/299,636
; FILING DATE: 02-SEP-1994
; CLASSIFICATION: 534
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/092,911
; FILING DATE: 18-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16777/262/ALLE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 1
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The Xaa at position 1 = sulphur"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 2
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The Xaa at position 2 = Acn"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 3
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The Xaa at position 3 = Mercaptoacetyl"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 5
; OTHER INFORMATION: /product= "OTHER"

```
; OTHER INFORMATION: /note= "The Xaa at position 5 = nitrogen"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The Xaa at position 6 = methylhydrazino nicotinic acid"
US-08-299-636-39

Query Match      100.0%; Score 29; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 8 TKPPR 12

RESULT 41
US-09-381-546-13
; Sequence 13, Application US/09381546
; Patent No. 6451976
; GENERAL INFORMATION:
; APPLICANT: Trigen Limited
; TITLE OF INVENTION: BI- OR MULTIFUNCTIONAL MOLECULES BASED ON A DENDROASPIN
; FILE REFERENCE: P41007NO
; CURRENT APPLICATION NUMBER: US/09/381,546
; CURRENT FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: PCT/GB98/00848
; PRIOR FILING DATE: 1998-09-20
; PRIOR APPLICATION NUMBER: GB9705787.1
; PRIOR FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 13
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified
; OTHER INFORMATION: dendroaspin
US-09-381-546-13

Query Match      100.0%; Score 29; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 9 TKPPR 13

RESULT 42
US-09-252-991A-25421
; Sequence 25421, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25421
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25421

; OTHER INFORMATION: /note= "The Xaa at position 5 = nitrogen"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The Xaa at position 6 = methylhydrazino nicotinic acid"
US-08-299-636-39

Query Match      100.0%; Score 29; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 8 TKPPR 12

RESULT 41
US-09-381-546-13
; Sequence 13, Application US/09381546
; Patent No. 6451976
; GENERAL INFORMATION:
; APPLICANT: Trigen Limited
; TITLE OF INVENTION: BI- OR MULTIFUNCTIONAL MOLECULES BASED ON A DENDROASPIN
; FILE REFERENCE: P41007NO
; CURRENT APPLICATION NUMBER: US/09/381,546
; CURRENT FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: PCT/GB98/00848
; PRIOR FILING DATE: 1998-09-20
; PRIOR APPLICATION NUMBER: GB9705787.1
; PRIOR FILING DATE: 1997-03-20
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 13
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: modified
; OTHER INFORMATION: dendroaspin
US-09-381-546-13

Query Match      100.0%; Score 29; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 9 TKPPR 13

RESULT 42
US-09-252-991A-25421
; Sequence 25421, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25421
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25421

; OTHER INFORMATION: /note= "The Xaa at position 5 = nitrogen"
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 6
; OTHER INFORMATION: /product= "OTHER"
; OTHER INFORMATION: /note= "The Xaa at position 6 = methylhydrazino nicotinic acid"
US-08-299-636-39

Query Match      100.0%; Score 29; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 72 TKPPR 76

RESULT 43
US-09-252-991A-21156
; Sequence 21156, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21156
; LENGTH: 136
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21156

Query Match      100.0%; Score 29; DB 4; Length 136;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 86 TKPPR 90

RESULT 44
US-09-489-039A-8820
; Sequence 8820, Application US/09489039A
; Patent No. 8610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8820
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8820

Query Match      100.0%; Score 29; DB 4; Length 420;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 270 TKPPR 274

RESULT 45
US-09-367-206-1
; Sequence 1, Application US/09367206
```

; Patent No. 6326482
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: NSP Molecules
; FILE REFERENCE: P1223RIE
; CURRENT APPLICATION NUMBER: US/09/367,206
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/08847
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 60/082,767
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 60/113,296
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 1
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-367-206-1

Query Match 100.0%; Score 29; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 211 TKPPR 215

RESULT 46
US-09-367-206-21
; Sequence 21, Application US/09367206
; Patent No. 6326482
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: NSP Molecules
; CURRENT APPLICATION NUMBER: US/09/367,206
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/08847
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 60/082,767
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 60/113,296
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 21
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutation of SEQ ID NO:1
US-09-367-206-21

Query Match 100.0%; Score 29; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 211 TKPPR 215

RESULT 47
US-09-367-206-22
; Sequence 22, Application US/09367206
; Patent No. 6326482
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: NSP Molecules
; FILE REFERENCE: P1223RIE
; CURRENT APPLICATION NUMBER: US/09/367,206
; CURRENT FILING DATE: 1999-08-09

; PRIOR APPLICATION NUMBER: PCT/US99/08847
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 60/082,767
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 60/113,296
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 22
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutation of SEQ ID NO:1
US-09-367-206-22

Query Match 100.0%; Score 29; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 211 TKPPR 215

RESULT 48
US-09-367-206-23
; Sequence 23, Application US/09367206
; Patent No. 6326482
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; TITLE OF INVENTION: NSP Molecules
; FILE REFERENCE: P1223RIE
; CURRENT APPLICATION NUMBER: US/09/367,206
; CURRENT FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/08847
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: US 60/082,767
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: US 60/113,296
; PRIOR FILING DATE: 1998-12-22
; NUMBER OF SEQ ID NOS: 35
; SEQ ID NO 23
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Mutation of SEQ ID NO:1
US-09-367-206-23

Query Match 100.0%; Score 29; DB 4; Length 576;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 211 TKPPR 215

RESULT 49
US-09-857-076-42
; Sequence 42, Application US/09357076C
; Patent No. 6225120
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogs, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Kowsek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; TITLE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS

FILE REFERENCE: 00786/351001
CURRENT APPLICATION NUMBER: US/08/857,076C
CURRENT FILING DATE: 1997-05-15
NUMBER OF SEQ ID NOS: 114
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 42
LENGTH: 892
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-08-857-076-42

Query Match 100.0%; Score 29; DB 3; Length 892;
Best Local Similarity 100.0%; Pred. No. 9.7e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 81 TKPPR 85

RESULT 50
US-08-118-101A-2
Sequence 2, Application US/08118101A
Patent No. 5620892
GENERAL INFORMATION:
APPLICANT: Kurtz, Stephen E.
APPLICANT: Knickerbocker, Aron M.
APPLICANT: McCullough, John R.
TITLE OF INVENTION: A STRAIN OF SACHAROMYCES CEREVISIAE
TITLE OF INVENTION: EXPRESSING THE GENE ENCODING POTASSIUM TRANSPORTER MINK
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burton Rodney
STREET: P.O. Box 4000
CITY: Princeton
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08543-4000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,101A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gaul, Timothy J.
REGISTRATION NUMBER: 33,111
REFERENCE/DOCKET NUMBER: DC27
TELEPHONE: (609) 252-5901
TELEFAX: (609) 252-4526
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1235 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-118-101A-2

Query Match 100.0%; Score 29; DB 1; Length 1235;
Best Local Similarity 100.0%; Pred. No. 1.3e+03; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
Db 438 TKPPR 442

Search completed: March 3, 2004, 12:19:39
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 12:15:03 ; Search time 39 Seconds
(without alignments)
40.451 Million cell updates/sec

Title: US-09-871-974-2

Perfect score: 29

Sequence: 1 TKPPR 5

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database :

SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_mhc:*

9: sp_mhc:*

10: sp_plant:*

11: sp_rhodent:*

12: sp_virus:*

13: sp_virus:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriaph:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	73	017534	Q17534 caenorhabdi
2	29	100.0	106	Q29253	Q29253 sus scrofa
3	29	100.0	115	12 Q98676	Q98676 simian cyto
4	29	100.0	142	16 Q9PQ08	Q9PQ08 xyiella fas
5	29	100.0	142	16 Q87EU9	Q87EU9 xyiella fas
6	29	100.0	143	13 Q91996	Q91996 acipenser b
7	29	100.0	169	12 Q7T922	Q7T922 human adeno
8	29	100.0	190	15 Q8F4V4	Q8F4V4 leptospira
9	29	100.0	191	10 Q9M250	Q9M250 arabidopsis
10	29	100.0	198	12 Q98674	Q98674 simian cyto
11	29	100.0	212	10 Q8H435	Q8H435 oryza sativ
12	29	100.0	222	16 Q8U6C3	Q8U6C3 agrobacteri
13	29	100.0	226	5 Q45855	Q45855 caenorhabdi
14	29	100.0	227	11 Q810F4	Q810F4 rattus norv
15	29	100.0	227	11 Q7TML1	Q7TML1 mus musculu
16	29	100.0	256	10 Q8SA84	Q8SA84 zea mays (m

17	29	100.0	250	11 Q35455	Q35455 mus musculu
18	29	100.0	278	16 P74248	P74248 synechocyst
19	29	100.0	283	5 Q24859	Q24859 entamoeba h
20	29	100.0	316	4 Q60440	Q60440 homo sapien
21	29	100.0	325	10 Q9FLJ1	Q9FLJ1 arabidopsis
22	29	100.0	326	10 Q9FL70	Q9FL70 arabidopsis
23	29	100.0	334	11 Q8CAP5	Q8CAP5 mus musculu
24	29	100.0	334	11 Q8UUK5	Q8UUK5 agrobacteri
25	29	100.0	351	5 Q23287	Q23287 caenorhabdi
26	29	100.0	384	4 Q60439	Q60439 homo sapien
27	29	100.0	370	16 Q8ZG89	Q8ZG89 yersinia pe
28	29	100.0	375	10 Q80865	Q80865 arabidopsis
29	29	100.0	394	16 Q82NT5	Q82NT5 streptomyce
30	29	100.0	405	16 Q85JB7	Q85JB7 bradyrhizob
31	29	100.0	406	16 Q83A13	Q83A13 coxiella bu
32	29	100.0	409	11 Q9D9R4	Q9D9R4 mus musculu
33	29	100.0	409	11 Q8BHB7	Q8BHB7 mus musculu
34	29	100.0	410	2 Q45945	Q45945 coxiella bu
35	29	100.0	410	2 Q52878	Q52878 coxiella bu
36	29	100.0	415	10 Q9XIF9	Q9XIF9 arabidopsis
37	29	100.0	415	10 Q8S8E1	Q8S8E1 arabidopsis
38	29	100.0	415	17 Q95Z07	Q95Z07 sulfolobus
39	29	100.0	429	10 Q944S3	Q944S3 arabidopsis
40	29	100.0	432	11 Q9D2W5	Q9D2W5 mus musculu
41	29	100.0	442	3 Q06349	Q06349 saccharomyc
42	29	100.0	458	5 Q9BLQ2	Q9BLQ2 meloidogyne
43	29	100.0	460	10 Q94LM4	Q94LM4 oryza sativ
44	29	100.0	491	16 Q828Y7	Q828Y7 streptomyce
45	29	100.0	499	16 Q82GE0	Q82GE0 streptomyce
46	29	100.0	505	16 Q7UT23	Q7UT23 rhodopirell
47	29	100.0	534	11 Q9D318	Q9D318 mus musculu
48	29	100.0	544	11 Q9EQY1	Q9EQY1 rattus norv
49	29	100.0	547	5 Q9VQ69	Q9VQ69 drosophila
50	29	100.0	574	4 Q96HX2	Q96HX2 homo sapien
51	29	100.0	575	5 Q62223	Q62223 caenorhabdi
52	29	100.0	576	4 Q9Y2X4	Q9Y2X4 homo sapien
53	29	100.0	576	4 Q9BRG2	Q9BRG2 homo sapien
54	29	100.0	579	11 Q9EQY2	Q9EQY2 rattus norv
55	29	100.0	585	12 Q7IF18	Q7IF18 rhesus cyto
56	29	100.0	615	11 Q9EQY3	Q9EQY3 rattus norv
57	29	100.0	619	4 Q8NGF6	Q8NGF6 homo sapien
58	29	100.0	669	10 Q8GXJ4	Q8GXJ4 arabidopsis
59	29	100.0	678	16 Q8PKG1	Q8PKG1 xanthomonas
60	29	100.0	722	5 Q9XTS1	Q9XTS1 caenorhabdi
61	29	100.0	776	11 Q8BQ16	Q8BQ16 mus musculu
62	29	100.0	808	2 Q700Z1	Q700Z1 saccharopol
63	29	100.0	809	10 Q81737	Q81737 arabidopsis
64	29	100.0	809	10 Q94AB2	Q94AB2 arabidopsis
65	29	100.0	816	10 Q9M2F3	Q9M2F3 arabidopsis
66	29	100.0	832	3 Q871X8	Q871X8 neurospora
67	29	100.0	854	5 Q9U0T9	Q9U0T9 leishmania
68	29	100.0	866	13 Q8UHU7	Q8UHU7 brachydanio
69	29	100.0	892	5 Q95Q17	Q95Q17 caenorhabdi
70	29	100.0	906	4 Q8N4W1	Q8N4W1 homo sapien
71	29	100.0	941	10 Q9SMD9	Q9SMD9 arabidopsis
72	29	100.0	959	10 Q8LGM9	Q8LGM9 arabidopsis
73	29	100.0	962	10 Q23048	Q23048 arabidopsis
74	29	100.0	988	5 Q9N8K2	Q9N8K2 trypanosoma
75	29	100.0	1003	4 Q9HCH0	Q9HCH0 homo sapien
76	29	100.0	1003	16 Q98GJ5	Q98GJ5 rhizobium 1
77	29	100.0	1022	5 Q27779	Q27779 schistosoma
78	29	100.0	1047	4 Q60284	Q60284 homo sapien
79	29	100.0	1067	5 Q7YU69	Q7YU69 drosophila
80	29	100.0	1413	5 Q9VK08	Q9VK08 drosophila
81	29	100.0	1430	11 Q8VHK2	Q8VHK2 rattus norv
82	29	100.0	1491	5 Q68BT9	Q68BT9 giardia lam
83	29	100.0	1667	5 Q9VHT6	Q9VHT6 drosophila
84	29	100.0	1706	5 Q9VAF5	Q9VAF5 drosophila
85	26	89.7	33	2 Q7WR15	Q7WR15 escherichia
86	26	89.7	42	16 Q8KE21	Q8KE21 chlorobium
87	26	89.7	47	16 Q9JXF4	Q9JXF4 neisseria m
88	26	89.7	50	16 Q8YTS4	Q8YTS4 anabaena sp
89	26	89.7	65	2 Q7WRH8	Q7WRH8 escherichia

```

90 26 89.7 66 16 Q9JZ50
91 26 89.7 68 16 Q9S279
92 26 89.7 75 16 Q9ZB10
93 26 89.7 81 17 Q8ZWE1
94 26 89.7 86 6 Q9GME3
95 26 89.7 87 5 Q8XKP6
96 26 89.7 87 16 Q8ZE23
97 26 89.7 88 12 Q89405
98 26 89.7 89 10 Q7XNQ7
99 26 89.7 100 17 Q9YCU1
100 26 89.7 101 12 Q9JIA0
101 26 89.7 102 12 Q806X6
102 26 89.7 105 16 Q88KU2
103 26 89.7 110 10 Q8S2F3
104 26 89.7 117 16 Q7UUS8
105 26 89.7 123 2 Q7WX19
106 26 89.7 127 16 Q8DBG4
107 26 89.7 132 12 Q9YMF3
108 26 89.7 139 11 Q80ZS9
109 26 89.7 139 16 Q8KB15
110 26 89.7 140 5 Q9ULJ6
111 26 89.7 140 5 Q9YMP9
112 26 89.7 141 10 Q8S3S2
113 26 89.7 147 10 Q7X6P1
114 26 89.7 148 16 Q9A7H4
115 26 89.7 149 17 Q8TV08
116 26 89.7 153 15 Q86322
117 26 89.7 153 16 P74047
118 26 89.7 154 5 Q9NEH6
119 26 89.7 154 17 Q30100
120 26 89.7 156 16 Q8R8V4
121 26 89.7 157 3 Q87IH3
122 26 89.7 163 5 Q8MZM8
123 26 89.7 167 4 Q8WTX0
124 26 89.7 171 16 Q82N66
125 26 89.7 172 4 Q9NWD7
126 26 89.7 173 12 Q8E184
127 26 89.7 179 16 Q8YX74
128 26 89.7 183 10 Q7XRQ8
129 26 89.7 183 11 Q9CWA7
130 26 89.7 183 16 Q9ISQ0
131 26 89.7 185 5 Q61697
132 26 89.7 189 10 Q9ARV1
133 26 89.7 192 12 Q8B8T7
134 26 89.7 192 12 Q7T940
135 26 89.7 192 12 Q7T8D1
136 26 89.7 195 2 Q50902
137 26 89.7 196 16 Q87XW1
138 26 89.7 198 5 Q7YVI3
139 26 89.7 201 16 Q82SQ3
140 26 89.7 213 2 Q7X040
141 26 89.7 213 12 Q8QS79
142 26 89.7 218 2 Q844P8
143 26 89.7 218 2 Q844P7
144 26 89.7 218 2 Q844N6
145 26 89.7 220 12 Q92486
146 26 89.7 220 4 Q9UFE7
147 26 89.7 230 17 Q8ZV71
148 26 89.7 231 10 Q7XVZ1
149 26 89.7 231 12 Q8B6S1
150 26 89.7 232 17 Q9YA94

```

ALIGNMENTS

```

RESULT 1
ID O17534 PRELIMINARY; PRT; 73 AA.
AC O17534; (Created)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)

```

```

Q9JZ50 neisseria m
Q9S279 streptomyce
Q9ZB10 listeria in
Q8ZWE1 pyrobaculum
Q9GME3 bos taurus
Q8XKP6 drosophila
Q8ZE23 yersinia pe
Q89405 paramecium
Q7XNQ7 oryza sativ
Q9YCU1 aeropyrum p
Q9JIA0 suid herpes
Q806X6 hepatitis c
Q88KU2 pseudomonas
Q8S2F3 oryza sativ
Q7UUS8 rhodospirell
Q7WX19 alcaligenes
Q8DBG4 vibrio vuln
Q9YMF3 orf virus.
Q80ZS9 mus musculu
Q8KB15 chlorobium
Q9ULJ6 caenorhabdi
Q9YMP9 caenorhabdi
Q8S3S2 physcomitre
Q7X6P1 oryza sativ
Q9A7H4 caulobacter
Q8TV08 methanopyru
Q86322 murine leuk
P74047 synchocyst
Q9NEH6 drosophila
Q30100 archaeoglob
Q8R8V4 thermoaer
Q87IH3 neurospora
Q8MZM8 taenia soli
Q8WTX0 homo sapien
Q82N66 streptomyce
Q9NWD7 homo sapien
Q8E184 san miguel
Q8YX74 anabaena sp
Q7XRQ8 oryza sativ
Q9CWA7 mus musculu
Q9ISQ0 pseudomonas
Q61697 manduca sex
Q9ARV1 oryza sativ
Q8B8T7 human adeno
Q7T940 human adeno
Q7T8D1 human adeno
Q50902 myxococcus
Q87XW1 pseudomonas
Q7YVI3 trypanosoma
Q82SQ3 nitrosomona
Q7X040 gamma-prote
Q8QS79 chimpanzee
Q844P8 gamma-prote
Q844P7 gamma-prote
Q844N6 gamma-prote
Q92486 bombyx mori
Q9UFE7 homo sapien
Q8ZV71 pyrobaculum
Q7XVZ1 oryza sativ
Q8B6S1 flanders vi
Q9YA94 aeropyrum p

```

```

DE DAF-3 (Fragment).
GN DAF-3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2;
RA Patterson G., Kowek A., Wong A., Liu Y., Ruvkun G.;
RT "The DAF-3 Smad protein antagonizes DAF-7 TGF-beta-receptor signalling
in the C. elegans dauer pathway.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF005207; AAB61750.1; -.
FT NON_TER 1 1
FT NON_TER 73 73
SQ SEQUENCE 73 AA; 8061 MW; 7D854125D687F0EE CRC64;

```

```

Query Match 100.0%; Score 29; DB 5; Length 73;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TKPR 5
Db 15 TKPR 19

```

RESULT 2

```

ID Q29253 PRELIMINARY; PRT; 106 AA.
AC Q29253;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Cytochrome C oxidase polypeptide III (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OC NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Small intestine;
RX MEDLINE=96327607; PubMed=8672129;
RA Winteroe A.K., Fredholm M., Davies W.;
RT "Evaluation and characterization of a porcine small intestine cDNA
library.";
RL Mamm. Genome 7:509-517(1996).
DR EMBL; F14871; CAA23306.1; -.
FT NON_TER 1 1
FT NON_TER 106 106
SQ SEQUENCE 106 AA; 12552 MW; 4BF91EA407F2EP73 CRC64;

```

```

Query Match 100.0%; Score 29; DB 6; Length 106;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TKPR 5
Db 48 TKPR 52

```

RESULT 3

```

ID Q98676 PRELIMINARY; PRT; 115 AA.
AC Q98676;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE BRLF3 (Fragment).
OS Simian cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OC NCBI_TaxID=10364;

```

RN SEQUENCE FROM N.A.
 RP PubMed:11725047;
 RA Chang Y., Jeang K., Lietman T., Hayward G.S.;
 RT "Structural organization of the Spliced Immediate-Early Gene Complex
 RT that Encodes the Major Acidic Nuclear (IE1) and Transactivator (IE2)
 RT Proteins of African Green Monkey Cytomegalovirus.";
 RL J. Biol. Med. Sci. 2:105-130(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90080130; PubMed=2152815;
 RA Chang Y.N., Crawford S., Stall J., Rawlins D.R., Jeang K.T.,
 RA Hayward G.S.;
 RT "The palindromic series I repeats in the simian cytomegalovirus major
 RT immediate-early promoter behave as both strong basal enhancers and
 RT cyclic AMP response elements.";
 RL J. Virol. 64:264-277(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=8719884; PubMed=3033283;
 RA Jeang K.T., Rawlins D.R., Rosenfield P.J., Shero J.H., Kelly T.J.,
 RA Hayward G.S.;
 RT "Multiple tandemly repeated binding sites for cellular nuclear factor
 RT 1 that surround the major immediate-early promoters of simian and
 RT human cytomegalovirus.";
 RL J. Virol. 61:1559-1570(1987).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93100836; PubMed=8380090;
 RA Chang Y.N., Jeang K.T., Chiu C.J., Chan Y.J., Pizzorno M.,
 RA Hayward G.S.;
 RT "Identification of a large bent DNA domain and binding sites for serum
 RT response factor adjacent to the NF1 repeat cluster and enhancer region
 RT in the major IE94 promoter from simian cytomegalovirus.";
 RL J. Virol. 67:516-529(1993).
 DR EMBL; U38308; A016875.1; -.
 FT NON TER 1
 SQ SEQUENCE 115 AA; 12965 MW; C4251671PBD87AF2 CRC64;

Query Match 100.0%; Score 29; DB 12; Length 115;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 110 TKPPR 114

RESULT 4
 Q9PGQ8

ID Q9PGQ8 PRELIMINARY; PRT; 142 AA.
 AC Q9PGQ8;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical protein Xf0240.
 GN Xf0240.
 OS Xylella fastidiosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=2371;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=9a5C;
 RX MEDLINE=20365717; PubMed=10910347;
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
 RA Alvarenga R., Alves L.M.C., Araya J.E., Bata G.S., Baptista C.S.,
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.N.,
 RA Coutinho L.B., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
 RA Fraga J.S., Franco S.C., Franco M.C., Frohme M., Furian L.R.,

RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
 RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
 RA Lemos E.G.M., Lemos M.V.P., Lopes S.A., Lopes C.R., Machado J.A.,
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
 RA Quaggio R.B., Roberto P.G., Rodrigues V. de Rosa A.J.M.,
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
 RA da Silva A.C.R., da Silva A.M., da Silva P.R., Silva W.A. Jr.,
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tshako M.H.,
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
 RL Nature 406:151-159(2000).
 DR EMBL; AE003878; AAF83053.1; -.
 DR PIR; C82829; C82829.
 DR InterPro; IPR002577; DUF24.
 DR Pfam; PF01638; DUF24; 1.
 DR ProDom; PD004032; DUF24; 1.
 DR Hypothetical protein; Complete proteome.
 KW SEQUENCE 142 AA; 15878 MW; CAD3A2B6EDB13E31 CRC64;

Query Match 100.0%; Score 29; DB 16; Length 142;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 87 TKPPR 91

RESULT 5
 Q87EU9

ID Q87EU9 PRELIMINARY; PRT; 142 AA.
 AC Q87EU9;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Conserved hypothetical protein.
 GN PD0199.
 OS Xylella fastidiosa (strain Temecul / ATCC 700964).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xylella.
 OX NCBI_TaxID=183190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22421331; PubMed=12533478;
 RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B.,
 RA Miyaki C.Y., Furian L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
 RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva P.R.,
 RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorri H., Tsai S.M.,
 RA Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,
 RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
 RA Marino C.L., Gigliotti E., Abreu I.L., Alves L.M.C., do Amaral A.M.,
 RA Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
 RA da Cunha A.F., Fenille R.C., Brito J.A., Formighieri E.F., Kishi L.T.,
 RA Leoni S.G., Oliveira A.R., Rosa V.E. Jr., Sasaki F.T., Sena J.A.D.,
 RA de Souza A.A., Truffi D., Tshako M.H., Yanai G.M., Zaros L.G.,
 RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
 RA Kitajima J.P.;
 RT "Comparative analyses of the complete genome sequences of Pierce's
 RT disease and citrus variegated chlorosis strains of Xylella
 RT fastidiosa.";
 RL J. Bacteriol. 195:1018-1026(2003).
 DR EMBL; AE012553; AAO28090.1; -.
 DR InterPro; IPR002577; DUF24.
 DR Pfam; PF01638; DUF24; 1.

KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 142 AA; 15808 MW; D52ED9B736C13E31 CRC64;

Query Match 100.0%; Score 29; DB 16; Length 142;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
| | | | |
DB 87 TKPPR 91

RESULT 6
Q91996 PRELIMINARY; PRT; 143 AA.
AC Q91996;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Thymoid-stimulating hormone precursor.
GN TSH.
OS Acipenser baerii (Siberian sturgeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Chondrostei; Acipenseriformes; Acipenseridae;
OC Acipenser.
OX NCBI_TaxID=27699;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=pituitary;
RX MEDLINE=20318422; PubMed=10859263;
RA Querat B., Sellouk A., Salmon C.;
RT "Phylogenetic analysis of the vertebrate glycoprotein hormone family including new sequences of sturgeon (Acipenser baerii) subunits of the two gonadotropins and the thyroid stimulating hormone.";
RL Biol. Reprod. 63:222-228(2000).
DR EMBL; AJ251659; CAB93505.1; -;
DR HSPSP; P01233; 1HCN.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR006208; Cys knot.
DR InterPro; IPR001545; Gly_hormoneB.
DR Pfam; PF00007; Cys knot; 1.
DR SMART; SM00068; GHb; 1.
DR PROSITE; PS00261; GLYCO_HORMONE_BETA_1; 1.
DR PROSITE; PS00689; GLYCO_HORMONE_BETA_2; 1.
KW Signal.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 143 THYROID-STIMULATING HORMONE.
SQ SEQUENCE 143 AA; 15566 MW; B79E009F7F0ED315 CRC64;

Query Match 100.0%; Score 29; DB 13; Length 143;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
| | | | |
DB 126 TKPPR 130

RESULT 7
Q7T922 PRELIMINARY; PRT; 169 AA.
AC Q7T922;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Agnoprotein.
GN Human adenovirus B.
OS Viruses; dsDNA viruses, no RNA stage; Adenoviridae; Mastadenovirus.
OX NCBI_TaxID=108098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=35p;

RX PubMed=12857895;
RA Vogels R., Zuidgeest D., Van Rijnsoever R., Hartkoorn E., Damen I.,
RA De Bethune M.P., Kostense S., Penders G., Helmus N., Koudstaal W.,
RA Cecchini M., Weterwald A., Sprangers M., Lemckert A., Ophorst O.,
RA Koel B., Van Meerendonk M., Quax P., Panitti L., Grimbergen J.,
RA Bout A., Goudsmit J., Havenga M.;
RT "Replication-Deficient Human Adenovirus Type 35 Vectors for Gene Transfer and Vaccination: Efficient Human Cell Infection and Bypass of Preexisting Adenovirus Immunity.";
RL J. Virol. 77:8263-8271(2003).
DR EMBL; AY271307; AAP92368.1; -;
SQ SEQUENCE 169 AA; 18824 MW; DC18351F66B90119 CRC64;

Query Match 100.0%; Score 29; DB 12; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
| | | | |
DB 97 TKPPR 101

RESULT 8
Q8F4V4 PRELIMINARY; PRT; 130 AA.
AC Q8F4V4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Conserved hypothetical protein with UVR motif.
GN LA1935.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB011366; AAN49134.1; -;
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0004518; F:nuclease activity; IEA.
DR GO; GO:0006289; P:nucleotide-excision repair; IEA.
DR InterPro; IPR003729; DUF151.
DR InterPro; IPR001943; UvrB/C.
DR Pfam; PF02577; DUF151; 1.
DR Pfam; PF02151; UVR; 1.
DR PROSITE; PS0151; UVR; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 130 AA; 21142 MW; 9A135B59D646E0AD CRC64;

Query Match 100.0%; Score 29; DB 16; Length 190;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
| | | | |
DB 52 TKPPR 56

RESULT 9
Q9M250 PRELIMINARY; PRT; 191 AA.
AC Q9M250;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN F7M19 120.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nyakatura G., Fartmann B., Dauner D., Sterr W., Holland R.,
 RL Weichselgartner M., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.,
 RA Quetier F., Salanoubat M.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL138643; CAB86482.1; -;
 DR PIR; T47369; T47369.
 KW Hypothetical protein.
 SQ SEQUENCE 191 AA; 21741 MW; 26F8764BEECB8A85D CRC64;
 Query Match 100.0%; Score 29; DB 10; Length 191;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TKPPR 5
 Db 78 TKPPR 82
 RESULT 10
 Q98674 PRELIMINARY; PRT; 198 AA.
 ID Q98674
 AC Q98674;
 DT 01-FEB-1997 (TREMBLrel. 02, Created)
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE UORF3.
 OS Simian cytomegalovirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae; Cytomegalovirus.
 OX NCBI_TaxID=10364;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=11725047;
 RA Chang Y., Jeang K., Lietman T., Hayward G.S.;
 RT "Structural Organization of the Spliced Immediate-Early Gene Complex that Encodes the Major Acidic Nucleat (IE1) and Transactivator (IE2) Proteins of African Green Monkey Cytomegalovirus.";
 RL J. Biomed. Sci. 2:105-130 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90080130; PubMed=2152815;
 RA Chang Y.N., Crawford S., Stall J., Rawlins D.R., Jeang K.T.,
 RA Hayward G.S.;
 RT "The palindromic series I repeats in the simian cytomegalovirus major immediate-early promoter behave as both strong basal enhancers and cyclic AMP response elements.";
 RL J. Virol. 64:264-277 (1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87198884; PubMed=3033283;
 RA Jeang K.T., Rawlins D.R., Rosenfeld P.J., Shero J.H., Kelly T.J.,
 RA Hayward G.S.;
 RT "Multiple tandemly repeated binding sites for cellular nuclear factor 1 that surround the major immediate-early promoters of simian and human cytomegalovirus.";
 RL J. Virol. 61:1559-1570 (1987).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93100836; PubMed=9380090;
 RA Chang Y.N., Jeang K.T., Chiou C.J., Chan Y.J., Pizzorno M.,
 RA Hayward G.S.;
 RT "Identification of a large bent DNA domain and binding sites for serum response factor adjacent to the NF1 repeat cluster and enhancer region in the major IE84 promoter from simian cytomegalovirus.";
 RL J. Virol. 67:516-529 (1993).
 DR EMBL; U38308; AB16873.1; -;

SQ SEQUENCE 198 AA; 22221 MW; DB6CE552D3775B0A CRC64;
 Query Match 100.0%; Score 29; DB 12; Length 198;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TKPPR 5
 Db 193 TKPPR 197
 RESULT 11
 Q98435 PRELIMINARY; PRT; 212 AA.
 ID Q98435
 AC Q98435;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE P0407H12.35 protein.
 GN P0407H12.35.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, PAC clone: P0407H12.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP004303; BAC21456.1; -;
 SQ SEQUENCE 212 AA; 24492 MW; 0EB7FD1860C63994 CRC64;
 Query Match 100.0%; Score 29; DB 10; Length 212;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TKPPR 5
 Db 108 TKPPR 112
 RESULT 12
 Q8U6C3 PRELIMINARY; PRT; 222 AA.
 ID Q8U6C3
 AC Q8U6C3;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein Atu4884.
 GN ATU4884 OR AGR_L_19.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608550; PubMed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eissen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58.";
 RL Science 294:2317-2323 (2001).

```

RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=21608551; PubMed=11743184; Miller N., Blanchard M., Mullin L.,
RA  Goodner B., Hinkle G., Gattung S., Cao Y., Askenazi M., Halling C., Liu P.,
RA  Houlihan K., Gordon J., Vaudin M., Tarchouk O., Epp A., Liu P.,
RA  Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA  Planagan C., Crowell C., Gursion J., Lomo C., Strub G.,
RA  Cielo C., Slater S.;
RT  "Genome sequence of the plant pathogen and biotechnology agent
RT  Agrobacterium tumefaciens C58."
RL  Science 294:2323-2328(2001).
DR  EMBL; AB009416; AAL45678.1; ALT_INIT.
DR  EMBL; AB008198; AAK88561.1; -.
DR  PIR; AH3157; AH3157.
DR  PIR; G98129; G98129.
DR  GO; GO:0003677; F:DNA binding; IEA.
DR  GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR  InterPro; IPR002145; HTH_CoPG.
DR  Pfam; PF01402; HTH_4; 1.
RW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 222 AA; 24796 MW; C8E8A5C3043B6EF3 CRC64;

Query Match 100.0%; Score 29; DB 16; Length 222;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 77 TKPPR 81

RESULT 13
C45855
ID O45855 PRELIMINARY; PRT; 226 AA.
AC O45855;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE T2E7.1 protein.
GN T2E7.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Cummings P.N.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
DR EMBL; Z82284; CAB05287.1; -.
DR PIR; T25360; T25360.
DR WormPep; T2E7.1; CE16505.
SQ SEQUENCE 226 AA; 25159 MW; 1C98B5A2873B6737 CRC64;

Query Match 100.0%; Score 29; DB 5; Length 226;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 75 TKPPR 79

RESULT 14
Q810F4 PRELIMINARY; PRT; 227 AA.
ID Q810F4

```

```

AC Q810F4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE FAM3C-like protein.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Bone marrow;
RA Buki K.G., Vaananen K.;
RT "Novel genes in rat bone marrow."
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY228475; AAO73558.1; -.
SQ SEQUENCE 227 AA; 24713 MW; 43A84E5B326D44A CRC64;

Query Match 100.0%; Score 29; DB 11; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 51 TKPPR 55

RESULT 15
Q7TML1 PRELIMINARY; PRT; 227 AA.
ID Q7TML1
AC Q7TML1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE D6Wsl76e protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Breast tumor;
RX MEDLINE=22389257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC055853; AAH55853.1; -.
SQ SEQUENCE 227 AA; 24783 MW; 98F02167EBB68CE1 CRC64;

Query Match 100.0%; Score 29; DB 11; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;

```

```
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
Db 51 TKPPR 55

RESULT 16
Q85A84 PRELIMINARY; PRT; 256 AA.
AC Q85A84;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Z195D10.20 protein (Fragment).
GN Z195D10.20.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. B73;
RA Ramakrishna W., Emberton J., SanMiguel P., Bennerzen J.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. B73;
RA Doesley J.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. B73;
RA Liaca V., Linton E.W., Young S., Kovchok S., Messing J.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF466646; AAL76008.1; -.
DR InterPro; IPR005333; TCP.
DR Pfam; PF03634; TCP; 1.
FT NON TER 256
SQ SEQUENCE 256 AA; 27971 MW; 79596CC55483648E CRC64;

Query Match 100.0%; Score 29; DB 10; Length 256;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
Db 216 TKPPR 220

RESULT 17
O35455 PRELIMINARY; PRT; 260 AA.
AC O35455;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Homeobox protein Nkx2.6 (Fragment).
GN Nkx2-6 OR Nkx2.6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98145949; PubMed=9486544;
RA Nikolova M., Chen X., Lufkin T.;
RT "Nkx2.6 expression is transiently and specifically restricted to the
RT branchial region of pharyngeal-stage mouse embryos.";
RL Mech. Dev. 69:215-218 (1997).
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL; AF030113; AAB86406.1; -.
```

```
DR HSSP; P23441; 1FTT.
DR MGD; MGI-97351; Nkx2-6.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; F:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001356; Homeobox.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PSS0071; HOMEBOX_2; 1.
KW DNA-binding; Homeobox; Nuclear protein.
FT NON TER 1
SQ SEQUENCE 260 AA; 28360 MW; 3264612FA98EBF9B CRC64;

Query Match 100.0%; Score 29; DB 11; Length 260;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
Db 161 TKPPR 165

RESULT 18
P74248 PRELIMINARY; PRT; 278 AA.
AC P74248;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein slr1169.
GN SLR1169.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hikosawa M., Sugita M., Saito K., Kimura T.,
RA Hoshino T., Murakami A., Murakami A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136 (1996).
DR EMBL; D90913; BAA18342.1; -.
DR PIR; S75883; S75883.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 278 AA; 30966 MW; E6503E98DB8C81EA CRC64;

Query Match 100.0%; Score 29; DB 16; Length 278;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
Db 240 TKPPR 244

RESULT 19
Q24859 PRELIMINARY; PRT; 283 AA.
AC Q24859;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
```

```
OX NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HM-1:IMSS;
RX MEDLINE=95327678; PubMed=7604025;
RA Clark C.G., Roger A.J.;
RT "Direct evidence for secondary loss of mitochondria in Entamoeba
histolytica";
RL Proc. Natl. Acad. Sci. U.S.A. 92:6518-6521(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HM-1:IMSS;
RA Clark C.G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: L39933; AAC41578.1; -
DR PIR: T18299; T18299.
KW Hypothetical protein.
FT NON_TER 283
SQ SEQUENCE 283 AA; 32222 MW; DD43E869752F9697 CRC64;

Query Match 100.0%; Score 29; DB 5; Length 283;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 208 TKPPR 212

RESULT 20
O60440 PRELIMINARY; PRT; 316 AA.
ID O60440
AC O60440;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Alpha-actinin-2 associated LIM protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Xia H., Bredt D.S.;
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 LIM DOMAIN. THE LIM DOMAIN BINDS 2 ZINC
IONS.
CC -1- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR EMBL: AF002282; AAC16672.1; -
DR GO: GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro: IPR001781; LIM.
DR InterPro: IPR001478; PDZ.
DR InterPro: IPR006643; ZASP.
DR Pfam: PF00442; LIM; 1.
DR Pfam: PF00595; PDZ; 1.
DR ProDom: PD000094; LIM; 1.
DR SMART: SM00132; LIM; 1.
DR SMART: SM00228; PDZ; 1.
DR SMART: SM00735; ZM; 1.
DR PROSITE: PS50023; LIM_DOMAIN_2; 1.
DR PROSITE: PS50106; PDZ; 1.
DR LIM domain; Metal-binding; Zinc.
KW LIM domain; Metal-binding; Zinc.
SQ SEQUENCE 316 AA; 34266 MW; C6D0A5E8A5DA3FF90 CRC64;

Query Match 100.0%; Score 29; DB 4; Length 316;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 301 TKPPR 305
```

```
RESULT 21
Q9FLJ1 PRELIMINARY; PRT; 325 AA.
ID Q9FLJ1
AC Q9FLJ1;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Genomic DNA, chromosome 5, P1 clone:MIO24.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98290546; PubMed=9628582;
RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
Sequence features of the regions of 1,456,315 bp covered by nineteen
physically assigned P1 and TAC clones.";
RL DNA Res. 5:41-54(1998).
DR EMBL: AB010074; BAB11237.1; -
SQ SEQUENCE 325 AA; 35789 MW; 9887D8FDFAB1C6AA CRC64;

Query Match 100.0%; Score 29; DB 10; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 9 TKPPR 13

RESULT 22
Q9FL70 PRELIMINARY; PRT; 326 AA.
ID Q9FL70
AC Q9FL70;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cotton fiber expressed protein 1-like protein.
GN AT5G54300.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=98344145; PubMed=9679202;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
features of the regions of 1,381,565 bp covered by twenty one
physically assigned P1 and TAC clones.";
RL DNA Res. 5:131-145(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
Sakurai T., Satou M., Seki M., Shim P., Yamada K., Shinozaki K.,
Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
```


RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
 RA Ecker J., Theologis A., Davis R.W.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB010695; BAB10753.1; -;
 DR EMBL; AY098819; AAM20670.1; -;
 DR EMBL; BT000327; AAN15646.1; -;
 DR InterPro; IPR008480; DUF761.
 DR Pfam; PF05553; DUF761.1;
 SQ SEQUENCE 326 AA; 36448 MW; 65E005C2D2D23B1F5F CRC64;

Query Match 100.0%; Score 29; DB 10; Length 326;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 210 TKPPR 214

RESULT 23
 Q8CAP5 PRELIMINARY; PRT; 334 AA.
 AC Q8CAP5;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Hypothetical GRAM domain containing protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium.
 RA "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK038289; BAC29960.1; -;
 DR InterPro; IPR004182; GRAM_dom.
 DR Pfam; PF02893; GRAM; 1.
 DR SMART; SM00568; GRAM; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 334 AA; 37116 MW; 57AE3045256D81A9 CRC64;

Query Match 100.0%; Score 29; DB 11; Length 334;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 306 TKPPR 310

RESULT 24
 Q8UKK5 PRELIMINARY; PRT; 350 AA.
 AC Q8UKK5;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE ATP-dependent DNA ligase.
 GN ATU5097 OR AGR PAT 142.
 GN Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Plasmid AT.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
 OX NCBI_TaxID=176299;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura Y.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
 RA Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C.,
 RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RA "The genome of the natural genetic engineer Agrobacterium tumefaciens
 RT C58.";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Mullin L.,
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmlel K., Gordon J., Vaudin M., Tarchouk O., Epp A., Liu F.,
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RA "Genome sequence of the plant pathogen and biotechnology agent
 RT Agrobacterium tumefaciens C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL; AE008934; AAL45787.1; -;
 DR EMBL; AE007882; AAK90473.1; -;
 DR PIR; AE3171; AE3171.
 DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0003910; F:DNA ligase (ATP) activity; IEA.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR GO; GO:0006281; P:DNA repair; IEA.
 DR GO; GO:0006260; P:DNA replication; IEA.
 DR InterPro; IPR000977; DNA_ligase.
 DR Pfam; PF01068; DNA_ligase; 1.
 DR Pfam; PF04679; DNA_ligase_A_C; 1.
 DR PROSITE; PS00160; DNA_LIGASE_A3; 1.
 KW Ligase; Plasmid; Complete proteome.
 SQ SEQUENCE 350 AA; 39363 MW; 32C15E7081D8233E CRC64;

Query Match 100.0%; Score 29; DB 16; Length 350;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 2 TKPPR 6

RESULT 25
 Q23287 PRELIMINARY; PRT; 351 AA.
 AC Q23287;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein (PIP-1).
 GN ZC404.8 OR PIP-1.
 GN Caenorhabditis elegans.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
 OC Rhabditidae; Peloderinae; Caenorhabditis.
 OX NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99089613; PubMed=9851916;
 RA None;
 RA "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).

RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Bentley D., Le T.T.;
 RT "The sequence of C. elegans cosmid 2C404.";
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT "Direct Submission.";
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Ogura K., Kishimoto N., Mitani S., Gengyo-Ando K., Kohara Y.;
 RT "Translational control of maternal gip-1 mRNA by POS-1 and its
 interacting protein SPN-4 in Caenorhabditis elegans.";
 RL Development 0:0-0(2003).
 DR EMBL; U53363; AAA97963.1; .
 DR EMBL; AB052819; BAC65239.1; .
 DR PIR; T29369; T29369.
 DR HSSP; P09651; 1HAI.
 DR WormPep; ZC404.8; CE07558.
 DR GO; GO:0003676; F:nuclieic acid binding; IEA.
 DR InterPro; IPR000504; RNA_rec_mot.
 DR Pfam; PF00076; rtm; 1.
 DR SMART; SM00360; RRM; 1.
 DR PROSITE; PS0102; RRM; 1.
 DR PROSITE; PS00030; RRM_RNP_1; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 351 AA; 39213 MW; A83B7159472851E CRC64;
 Query Match 100.0%; Score 29; DB 5; Length 351;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 199 TKPPR 203

RESULT 26
 O60439 PRELIMINARY; PRT; 364 AA.
 ID O60439
 AC O60439
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Alpha-actinin-2 associated LIM protein.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RA Xia H., Bredt D.S.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 LIM DOMAIN. THE LIM DOMAIN BINDS 2 ZINC
 IONS.
 CC -1- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
 DR EMBL; AF002280; AAC16670.1; .
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
 DR InterPro; IPR001781; LIM.
 DR InterPro; IPR001478; PDZ.
 DR InterPro; IPR006643; ZASP.
 DR Pfam; PF00412; LIM; 1.
 DR Pfam; PF00595; PDZ; 1.
 DR ProDom; PD000094; LIM; 1.
 DR SMART; SM00132; LIM; 1.
 DR SMART; SM00228; PDZ; 1.
 DR SMART; SM00735; ZM; 1.
 DR PROSITE; PS50023; LIM_DOMAIN_2; 1.

DR PROSITE; PS50106; PDZ; 1.
 KW LIM domain; Metal-binding; Zinc.
 SQ SEQUENCE 364 AA; 39231 MW; EAAA2A75466D3E12 CRC64;
 Query Match 100.0%; Score 29; DB 4; Length 364;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 349 TKPPR 353

RESULT 27
 Q8ZG89 PRELIMINARY; PRT; 370 AA.
 ID Q8ZG89
 AC Q8ZG89
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Putative iron-sulfur binding protein (Hypothetical protein).
 GN YPO1417 OR Y2752.
 OS Yersinia pestis.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Yersinia.
 CX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Brooks K., Cerdano-Tarraga A.M.,
 RA Baker S., Baeham D., Bentley S.D., Davies R.M., Davis P., Dougan G.,
 RA Chillingworth T., Cronin A., Cronin A., Davis R.M., Dougan G.,
 RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leavelle S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIMS / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner P.R.,
 RA Perry R.D.;
 RT "Genome sequence of Yersinia pestis KIM.";
 RL J. Bacteriol. 184:4601-4611(2002).
 DR EMBL; AJ414148; CAC90246.1; .
 DR EMBL; AE013878; RAM86304.1; .
 DR PIR; AC0173; AC0173.
 DR GO; GO:0005489; F:electron transporter activity; IEA.
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR006058; 2Fe2S fd BS.
 DR InterPro; IPR001041; Ferredoxin.
 DR InterPro; IPR005302; MOSC.
 DR InterPro; IPR00111; fer2; 1.
 DR Pfam; PF00111; fer2; 1.
 DR Pfam; PF03473; MOSC; 1.
 DR Pfam; PF03476; MOSC_N; 1.
 DR PROSITE; PS00197; 2Fe2S FERREDOXIN; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 370 AA; 40747 MW; CF8246C4DC713138 CRC64;
 Query Match 100.0%; Score 29; DB 16; Length 370;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 266 TKPPR 270

DE Hypothetical protein.
GN SAV1206.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AF005025; BAC68916.1; -;
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR006090; Acyl-CoA dh_C.
DR Pfam; PF00441; Acyl-CoA dh; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 394 AA; 42238 MW; B8CF41D0C94F29D2 CRC64;

Query Match 100.0%; Score 29; DB 16; Length 394;
Best Local Similarity 100.0%; Pred.No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 386 TKPPR 390

RESULT 30
ID Q89JB7 PRELIMINARY; PRT; 405 AA.
AC Q89JB7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE B11366 protein.
GN B11366.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484398; PubMed=12597275;
RA Kaneko T., Nakamura F., Sato S., Minamisawa K., Uchiuni T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AF005954; BAC50631.1; -;
DR GO; GO:0004730; F:pseudouridyate synthase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR InterPro; IPR006145; PseudoU synth.
DR InterPro; IPR006224; Rlu synth.
DR InterPro; IPR002942; S4.

Query Match 100.0%; Score 29; DB 10; Length 375;
Best Local Similarity 100.0%; Pred.No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 121 TKPPR 125

RESULT 29
ID Q82NT5 PRELIMINARY; PRT; 394 AA.
AC Q82NT5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=16617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.B., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Sherr A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
RA Salberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=16617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.B., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Sherr A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
RA Salberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=16617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.B., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Sherr A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
RA Salberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=16617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.B., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Sherr A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
RA Salberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=16617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.B., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Sherr A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
RA Salberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=16617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.B., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Sherr A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
RA Salberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=16617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.B., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Sherr A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
RA Salberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=16617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.B., Feldblum T.V.,
RA Buehl C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Sherr A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nieman W.C., White O., Eisen J.A.,
RA Salberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]_
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487

!

```
Db 358 TKPPR 362

RESULT 34
ID Q45945 PRELIMINARY; PRT; 410 AA.
AC Q45945; Q45901;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE ORF 410.
OS Coccidia burnetii.
OG Plasmid QP41.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NINE MILE PHASE I;
RX MEDLINE=95145653; PubMed=7843345;
RA Thiele D., Willems H., Haas M., Krauss H.;
RT "Analysis of the entire nucleotide sequence of the cryptic plasmid
RT QpH1 from Coxiella burnetii.";
RL Eur. J. Epidemiol. 10:413-420(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NINE MILE PHASE I;
RA Thiele D.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SCURRY Q217;
RA Rittner M., Thiele D., Willems H.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; X75356; CAA53126.1; -.
DR EMBL; X93204; CAA63678.1; -.
DR PIR; S38238; S38238.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0015074; P:DNA recombination; IEA.
DR GO; GO:0006310; P:DNA recombination; IEA.
DR InterPro; IPR002104; Phage integrase.
DR Pfam; PF00589; Phage integrase; 1.
KW Plasmid.
SQ SEQUENCE 410 AA; 47829 MW; 9A48C3714B330C2 CRC64;

Query Match 100.0%; Score 29; DB 2; Length 410;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 102 TKPPR 106

RESULT 35
ID O52878 PRELIMINARY; PRT; 410 AA.
AC O52878;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Coccidia burnetii.
OG Plasmid QP41, and Plasmid QpDV.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Legionellales;
OC Coxiellaceae; Coxiella.
OX NCBI_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PRISCILLA Q177;
RA Lautenschlaeger S., Jaeger C., Willems H., Baljer G.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

Db 358 TKPPR 362

RESULT 36
ID Q9XIF9 PRELIMINARY; PRT; 415 AA.
AC Q9XIF9;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE T10P12.9 protein.
GN T10P12.9.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Alafifi H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,
RA Luros S., Schwartz J., Shinn P., Toriumi M., Vysotskaya V.S.,
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007203; AAD39275.1; -.
DR PIR; F96499; F96499.
DR HSSP; P50586; LC8Z.
DR InterPro; IPR000007; Tubby.
DR Pfam; PF01167; Tub; 1.
DR PRINTS; PR01573; SUPERTUBBY.
DR PROSITE; PS01200; TUB_1; 1.
SQ SEQUENCE 415 AA; 46492 MW; 38774871B95D8770 CRC64;

Query Match 100.0%; Score 29; DB 10; Length 415;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
Db 309 TKPPR 313

RESULT 37
ID Q8S8E1 PRELIMINARY; PRT; 415 AA.
AC Q8S8E1;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
```

Qy

```

RESULT 40
Q9D2W5 PRELIMINARY; PRT; 432 AA.
AC Q9D2W5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE 9130427A09Rik protein.
DE 9130427A09Rik.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Cecum;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukuishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gajobori T., Boro H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Mateuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baidarelli R., Barsh G.,
RA Blake J., Boicelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohnsuke S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK018691; BAB31347.1; -.
DR MGD; MGI:1914815; 9130427A09Rik.
DR InterPro; IPR004182; GRAM_dom.
DR Pfam; PF02893; GRAM; 1.
DR SMART; SM00568; GRAM; 1.
SQ SEQUENCE 432 AA; 47932 MW; 2D9898518F2B544B CRC64;

Query Match 100.0%; Score 29; DB 11; Length 432;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 300 TKPPR 304

RESULT 41
Q06349 PRELIMINARY; PRT; 442 AA.
AC Q06349;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Chromosome IV COSMID 9481.
GN YDR370C OR D9481.14.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RA Ding H.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

RC STRAIN=S288C;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favell A., Fulton L., Gattung S., Greco T., Kirsten J., Kucaba T.,
RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson D.,
RA Johnston L., Langston Y., Latreille P., Le T., Mardis E., Meneses S.,
RA Miller N., Nhan M., Pauley A., Peluso D., Rifkin L., Riles L.,
RA Taich A., Trevasxis E., Vignati D., Wilcox L., Wohlman P., Vaudin M.,
RA Wilson R., Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=S288C;
RA Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Jia Y., Cherry J.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U28373; AAB64806.1; -.
DR PIR; S61165; S61165.
DR SGD; S0002778; YDR370C.
SQ SEQUENCE 442 AA; 50492 MW; 504F7C6AA40B50B4 CRC64;

Query Match 100.0%; Score 29; DB 3; Length 442;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 81 TKPPR 85

RESULT 42
Q9BLQ2 PRELIMINARY; PRT; 458 AA.
AC Q9BLQ2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE Putative avirulence protein precursor.
GN MAP-1.
OS Meloidogyne incognita (southern root-knot nematode).
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
OC Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
OX NCBI_TaxID=6306;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21034633; PubMed=11194874;
RA Sembat J.P., Rosso M.N., Husey R.S., Abad P., Castagnone-Sereno P.;
RT "Molecular cloning of a cDNA encoding an amphid-secreted putative
RT avirulence protein from the root-knot nematode Meloidogyne
RT incognita";
RL Mol. Plant Microbe Interact. 14:72-79(2001).
DR EMBL; AJ278663; CAC27774.1; -.
DR InterPro; IPR002965; P-rich_extensn.
DR PRINTS; PR01217; PRICHEXTENS.
KW SIGNAL.
FT SIGNAL. 1 17 POTENTIAL.
FT CHAIN 18 458 PUTATIVE AVIRULENCE PROTEIN.
SQ SEQUENCE 458 AA; 48951 MW; 982AD87D0818815F CRC64;

Query Match 100.0%; Score 29; DB 5; Length 458;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 168 TKPPR 172

RESULT 43
Q94LM4

```

ID Q94LM4 PRELIMINARY; PRT; 460 AA.
AC Q94LM4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Unknown protein.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Shariatoidae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA King R.A., Frisch D., Presting G., Wood T., Yu Y., Soderlund C.,
RA Kim R.A., Rambo T., Henry D., Simmons J.,
RA "Rice Genomic Sequence".
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC079038; AAK5572.1; -.
DR Gramine; Q94LM4; -.
DR InterPro; IPR008938; ARM.
SQ SEQUENCE 460 AA; 50005 MW; 8CFCEC923197779 CRC64;
Query Match 100.0%; Score 29; DB 10; Length 460;
Best Local Similarity 100.0%; Pred. No. 3.8e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;
QY 1 TKPPR 5
DB 136 TKPPR 140
RESULT 44
ID Q828Y7 PRELIMINARY; PRT; 491 AA.
AC Q828Y7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative GTP-binding protein.
GN SAV6524.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites".
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis".
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005047; BAC74235.1; -.
DR GO; GO:0001602; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006910; F:transport; IEA.
DR InterPro; IPR004638; Eflux_EmrB.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_cr; 1.
DR TIGRfams; TIGR00711; eflux_EmrB; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Complete proteome.
SQ SEQUENCE 499 AA; 52225 MW; F580C79D5469F2CF CRC64;
Query Match 100.0%; Score 29; DB 16; Length 499;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
DB 482 TKPPR 486
RESULT 46

DR SMART; SMO0382; AAA; 2.
DR TIGRfams; TIGR00650; MG442; 2.
DR TIGRfams; TIGR00231; small_GTP; 2.
KW Complete proteome.
SQ SEQUENCE 491 AA; 53353 MW; 7BAD8E0449D14AC CRC64;
Query Match 100.0%; Score 29; DB 16; Length 491;
Best Local Similarity 100.0%; Pred. No. 4e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;
QY 1 TKPPR 5
DB 440 TKPPR 444
RESULT 45
ID Q82GE0 PRELIMINARY; PRT; 499 AA.
AC Q82GE0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transport integral membrane protein.
GN SAV3957.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.,
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites".
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis".
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP005037; BAC71669.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006910; F:transport; IEA.
DR InterPro; IPR004638; Eflux_EmrB.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; sugar_cr; 1.
DR TIGRfams; TIGR00711; eflux_EmrB; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Complete proteome.
SQ SEQUENCE 499 AA; 52225 MW; F580C79D5469F2CF CRC64;
Query Match 100.0%; Score 29; DB 16; Length 499;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
DB 482 TKPPR 486
RESULT 46

Q7UT23

ID Q7UT23 PRELIMINARY; PRT; 505 AA.
 AC Q7UT23;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE 4-alpha-glucanotransferase (EC 2.4.1.25).
 GN MALQ OR RB4161.
 OS Rhodospirillum rubrum.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.
 OX NCBI_TaxID=117;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1;
 RX MEDLINE=22735913; PubMed=12835416;
 RA Glocker F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
 RA Schlesner H., Amann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.
 strain 1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
 RL EMBL; BX294140; CAD73618.1; -.
 KW Glycosyltransferase; Transferase; Complete proteome.
 SQ SEQUENCE 505 AA; 57353 MW; 5470834CCAR15PFF CRC64;

Query Match 100.0%; Score 29; DB 16; Length 505;
 Best Local Similarity 100.0%; Pred. NO. 4.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 DB 121 TKPPR 125

RESULT 47

Q9D318 PRELIMINARY; PRT; 514 AA.
 AC Q9D318;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE 9030613F08Rik protein.
 GN 9030613F08Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Colon;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischnann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuda Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RL EMBL; AK018551; BAB31270.1; -.

MGD; MGI:1921790; 9030613F08Rik.
 DR InterPro; IPR004182; GRAM_dom.
 DR Pfam; PF02893; GRAM; 1.
 DR SMART; SM00568; GRAM; 1.
 SQ SEQUENCE 514 AA; 56478 MW; CSA504279D228B84 CRC64;

Query Match 100.0%; Score 29; DB 11; Length 514;
 Best Local Similarity 100.0%; Pred. NO. 4.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 DB 382 TKPPR 386

RESULT 48

Q9EQY1 PRELIMINARY; PRT; 524 AA.
 AC Q9EQY1;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Transcription factor Elf-1.
 GN ELF-1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21077473; PubMed=11210123;
 RA Nishiyama C., Takahashi K., Nishiyama M., Okumura K., Ra C.,
 RA Ohtake Y., Yokota T.;
 RT "Polymorphism of transcription factor Elf-1 affecting its regulatory
 function in transcription.";
 RL Biosci. Biotechnol. Biochem. 64:2601-2607(2000).
 DR EMBL; AB030217; BAB20035.1; -.
 DR HSSP; P28324; 1BC8.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR000416; Ets.
 DR InterPro; IPR002341; HSF_ETS.
 DR Pfam; PF00178; Ets; 1.
 DR PRINTS; PR00454; ETSDOMAIN.
 DR SMART; SM00413; ETS; 1.
 DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
 DR PROSITE; PS00345; ETS_DOMAIN_2; 1.
 DR PROSITE; PS00061; ETS_DOMAIN_3; 1.
 SQ SEQUENCE 524 AA; 56580 MW; 87834210FC7AAD1A CRC64;

Query Match 100.0%; Score 29; DB 11; Length 524;
 Best Local Similarity 100.0%; Pred. NO. 4.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 DB 98 TKPPR 92

RESULT 49

Q9VQ69 PRELIMINARY; PRT; 547 AA.
 AC Q9VQ69; Q961C7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE CG31672 protein (LD22825p).
 GN BEST-LD15963 OR CG4248 OR CG31672.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champagne M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Flossler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glouck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimball B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mileshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RN "The genome sequence of Drosophila melanogaster."
 Science 287:2185-2195(2000).
 [2]
 RP SEQUENCE FROM N.A.
 RA Celnikier S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busan D.A.,
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Faragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome."
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RA Misra S., Crosby M.A., Matthews B.E., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaninker J.S., Prochnik S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnikier S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smtutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome."
 RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 [4]
 RP SEQUENCE FROM N.A.

RA Adams M.D., Celnikier S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champagne M., Chavez C., Dotsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Faragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celnikier S.,
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003584; AAF51310.1; -;
 DR EMBL; AY051677; AAK93101.1; -;
 DR FlyBase; FBGN0028952; BEST:LD15963.
 SQ SEQUENCE 547 AA; 63439 MW; 74CFFD0A1BA4F195 CRC64;
 Query Match 100.0%; Score 29; DB 5; Length 547;
 Best Local Similarity 100.0%; Pred. No. 4.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 Db 205 TKPPR 209
 RESULT 50
 O96HX2 PRELIMINARY; PRT; 574 AA.
 ID Q96HX2
 AC Q96HX2
 DT 01-DEC-2001 (TRENBLrel. 19, Created)
 DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC007998; AAH07998.1; -;
 KW Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 574 AA; 60606 MW; E598819330ADFE64 CRC64;
 Query Match 100.0%; Score 29; DB 4; Length 574;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 Db 387 TKPPR 391
 Search completed: March 3, 2004, 12:18:34
 Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 12:11:12 ; Search time 55 Seconds
(without alignments)
25.686 Million cell updates/sec

Title: US-09-871-974-2

Perfect score: 29

Sequence: 1 TKPRP 5 ←

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 150 summaries

Database : A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003Bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29	100.0	5	2	Aaw11052 Leukocyte
2	29	100.0	5	2	Aaw31146 Amyloid p
3	29	100.0	5	2	Aay49845 Tuftsin r
4	29	100.0	5	5	Abb08442 Tuftsin r
5	29	100.0	5	5	Abb08442 Tuftsin r
6	29	100.0	5	7	Aam51906 Tertiary
7	29	100.0	5	7	Add10684 Tuftsin a
8	29	100.0	6	2	Aaw11053 Leukocyte
9	29	100.0	6	2	Aaw31148 Amyloid p
10	29	100.0	6	2	Aaw31147 Amyloid p
11	29	100.0	6	2	Aay49840 Tuftsin r
12	29	100.0	6	5	Abb08444 Tuftsin r
13	29	100.0	6	5	Aam51904 Tertiary
14	29	100.0	6	5	Aam51905 Tertiary
15	29	100.0	6	7	Add10689 Tuftsin a
16	29	100.0	6	7	Add10687 Tuftsin a
17	29	100.0	7	2	Aay49842 Tuftsin r
18	29	100.0	7	2	Aay49841 Tuftsin r
19	29	100.0	7	5	Abb08447 Tuftsin r
20	29	100.0	7	5	Abb08445 Linear an
21	29	100.0	7	7	Add10688 Cyclic Tu
22	29	100.0	7	7	Add10686 Tuftsin a
23	29	100.0	8	2	Aar88740 Tuftsin a
24	29	100.0	8	2	Aar76218 peptide-l
25	29	100.0	8	2	Aaw11058 Leukocyte

26	29	100.0	8	2	Aay49844 Tuftsin r
27	29	100.0	8	5	Abb08448 Tuftsin r
28	29	100.0	9	2	Aar85539 Metal che
29	29	100.0	9	2	Aar85535 Metal che
30	29	100.0	9	2	Aar88735 Tuftsin a
31	29	100.0	9	2	Aar88738 Tuftsin a
32	29	100.0	9	2	Aar88741 Tuftsin a
33	29	100.0	9	2	Aar76219 Peptide-l
34	29	100.0	9	2	Aaw11055 Leukocyte
35	29	100.0	9	2	Aaw11059 Leukocyte
36	29	100.0	9	2	Aaw11054 Leukocyte
37	29	100.0	9	2	Aaw11056 Leukocyte
38	29	100.0	9	2	Aaw11057 Leukocyte
39	29	100.0	9	2	Aaw03420 Peptide u
40	29	100.0	9	2	Aay23752 Peptide R
41	29	100.0	10	2	Aar88739 Tuftsin a
42	29	100.0	10	7	Ade81165 Tuftsin m
43	29	100.0	10	7	Ade81166 Tuftsin m
44	29	100.0	11	2	Aar88736 Tuftsin a
45	29	100.0	11	7	Add10691 Tuftsin a
46	29	100.0	11	7	Add10690 Tuftsin a
47	29	100.0	12	5	Abb08446 Tuftsin r
48	29	100.0	20	5	Abb08449 99mTc rad
49	29	100.0	53	5	Abp34343 Human ORF
50	29	100.0	65	4	Abg07007 Novel hum
51	29	100.0	99	4	Aau51890 Propionib
52	29	100.0	99	5	Abb84958 Human PRO
53	29	100.0	99	5	Abg34045 Human PRO
54	29	100.0	99	5	Abb95564 Human ang
55	29	100.0	99	6	Abm48409 Propionib
56	29	100.0	99	6	Ada01300 Human PRO
57	29	100.0	99	6	Ada43729 Human sec
58	29	100.0	99	6	Ada43497 Human sec
59	29	100.0	99	6	Ada01172 Human PRO
60	29	100.0	99	7	Ada01056 Human sec
61	29	100.0	99	7	Ada43613 Human sec
62	29	100.0	99	7	Ada06875 Human PRO
63	29	100.0	99	7	Ada08363 Novel hum
64	29	100.0	99	7	Abd99856 Human PRO
65	29	100.0	99	7	Abd86939 Human PRO
66	29	100.0	99	7	Abd66094 Human sec
67	29	100.0	99	7	Abd99772 Human PRO
68	29	100.0	99	7	Abd99427 Novel hum
69	29	100.0	99	7	Abd65978 Human sec
70	29	100.0	99	7	Adc23376 Human tra
71	29	100.0	99	7	Adc23376 Human tra
72	29	100.0	99	7	Adc23376 Human tra
73	29	100.0	99	7	Add10573 Human sec
74	29	100.0	99	7	Add10573 Human sec
75	29	100.0	99	7	Adc04896 Human PRO
76	29	100.0	99	7	Adc04896 Human PRO
77	29	100.0	99	7	Adc04896 Human PRO
78	29	100.0	99	7	Add88133 Human PRO
79	29	100.0	99	7	Add95428 Human sec
80	29	100.0	99	7	Add95428 Human sec
81	29	100.0	99	7	Add95428 Human sec
82	29	100.0	99	7	Add88249 Human PRO
83	29	100.0	99	7	Add90830 Human PRO
84	29	100.0	99	8	Ade51683 Human sec
85	29	100.0	99	8	Ade51799 Human sec
86	29	100.0	99	8	Ade37657 Human sec
87	29	100.0	99	8	Ade37657 Human sec
88	29	100.0	99	8	Add95312 Human sec
89	29	100.0	99	8	Ade38012 Human PRO
90	29	100.0	99	8	Ade76101 Human PRO
91	29	100.0	99	8	Ade39424 Human PRO
92	29	100.0	99	8	Ade04228 Human PRO
93	29	100.0	99	8	Ade39825 Human PRO
94	29	100.0	99	8	Ade19890 Human PRO
95	29	100.0	99	8	Ade77268 Human sec
96	29	100.0	99	8	Ade65376 Human PRO
97	29	100.0	99	8	Ade75985 Human PRO
98	29	100.0	99	8	Ade37896 Human PRO
99	29	100.0	99	8	Ade64506 Human PRO

99 29 100.0 99 8 ADE41534
100 29 100.0 99 8 ADE38841 Human PRO
101 29 100.0 99 8 ADE51915
102 29 100.0 99 8 ADE90946 Human sec
103 29 100.0 99 8 ADE38725 Human PRO
104 29 100.0 99 8 ADE37425 Human sec
105 29 100.0 99 8 ADE06242 Human PRO
106 29 100.0 99 8 ADE90101 Human sec
107 29 100.0 99 8 ADE38609 Human PRO
108 29 100.0 99 8 ADE39540 Human PRO
109 29 100.0 99 8 ADE89145 Human PRO
110 29 100.0 99 8 ADE88912 Human PRO
111 29 100.0 99 8 ADE19806 Human PRO
112 29 100.0 99 8 ADE77384 Human sec
113 29 100.0 99 8 ADE5260 Human PRO
114 29 100.0 99 8 ADE39308 Human PRO
115 29 100.0 99 8 ADE38493 Human sec
116 29 100.0 103 6 AAU42638
117 29 100.0 103 6 ABM39157 Propionib
118 29 100.0 107 3 AAY70535 Maize pla
119 29 100.0 136 4 ABG22236 Novel hum
120 29 100.0 149 4 ABG16691 Novel hum
121 29 100.0 198 4 AAU46263 Propionib
122 29 100.0 198 6 ABM42782 Propionib
123 29 100.0 216 4 ABG03775 Novel hum
124 29 100.0 227 5 ABG07008 Novel hum
125 29 100.0 253 4 ABG07008 Novel hum
126 29 100.0 304 4 ABG16575 Novel hum
127 29 100.0 309 5 ABG77237 Selected
128 29 100.0 309 5 ABG11108 Yeast sel
129 29 100.0 310 4 ABG06185 Novel hum
130 29 100.0 334 4 ABG17530 Novel hum
131 29 100.0 340 3 AAG43777 Arabidops
132 29 100.0 343 3 AAG43776 Arabidops
133 29 100.0 344 3 AAG43775 Arabidops
134 29 100.0 350 4 AAB50667 C. elegan
135 29 100.0 351 4 AAB50666 C. elegan
136 29 100.0 409 7 AAE34477 Murine AG
137 29 100.0 415 7 ADC77485 Sulfolobu
138 29 100.0 415 7 ADD93932 Sulfolobu
139 29 100.0 474 4 ABB59612 Drosophil
140 29 100.0 494 4 ABG03323 Novel hum
141 29 100.0 524 3 AAB03192 Rat Elf-1
142 29 100.0 530 4 ABG03809 Novel hum
143 29 100.0 530 4 ABG01390 Novel hum
144 29 100.0 563 4 ABG04846 Novel hum
145 29 100.0 563 4 ABG16158 Novel hum
146 29 100.0 576 2 AAY06477 Human tum
147 29 100.0 576 2 AAY49546 Human Nsp
148 29 100.0 576 2 AAY49545 Human Nsp
149 29 100.0 576 2 AAY49547 Human Nsp
150 29 100.0 576 2 AAY49541 Human PRO

ALIGNMENTS

RESULT 1
AAW11052
ID AAW11052 standard; peptide; 5 AA.

XX AC AAW11052;
DT 03-JUN-1997 (first entry)
XX

DE Leukocyte-targeting peptide used in diagnostic imaging.

XX Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
KW diagnostic imaging.

OS Synthetic.

XX WO9603427-A1.

XX 08-FEB-1996.
PD
XX
PF 28-APR-1995; 95WO-CA000249.
XX
XX 22-JUL-1994; 94US-00279155.
PR
XX (RESO-) RESOLUTION PHARM INC.
PA Pollak A, Goodbody A;
FI WPI; 1996-116994/12.
XX
XX New peptide derived radionuclide chelators and metal complexes - useful
PT for diagnostic imaging.
XX
FS Claim 12; Page 19; 30pp; English.
XX
CC AAW11052 is a peptide used for targeting agents to leukocytes. This
CC peptide can be coupled to a metal radionuclide chelator (structurally
CC formula given in the specification), labelled with a diagnostically
CC useful metal isotope, to form a peptide derived radionuclide chelator.
CC When the chelator is coupled to a targeting molecule and labelled with a
CC diagnostically useful metal, it can be used to detect pathological
CC conditions by diagnostic imaging. For example, AAW1052 targets the
CC chelator to leukocytes and are useful for the rapid imaging of sites of
CC local inflammation. Radionuclides used include 99mTc, 64Cu, 67Cu, 97Ru,
CC 105Rh, 109Pd, 186Re, 188Re, 198Au, 199Au, 203Pb, 212Pb and 212Bi. The
CC coupling of a targeting agent and radionuclide using a chelating agent is
CC an alternative to the direct labelling of targeting agents in which
CC radionuclides are typically bound at the more numerous low-affinity
CC sites, forming unstable complexes. The new conjugates give better
CC scintigraphic images in rat inflammation studies than known imaging
CC agents Ga-67, 99mTc-1gG, 111In-WBC and 99mTc-Nanocol. They image more
CC rapidly than the known agents and show superior biodistribution
XX
SQ Sequence 5 AA;
Query Match 100.0%; Score 29; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
DB 1 TKPPR 5
RESULT 2
AAW31146
ID AAW31146 standard; peptide; 5 AA.
XX AC AAW31146;
XX
DT 25-MAR-2003 (revised)
DT 23-JAN-1998 (first entry)
XX
DE Amyloid plaque-targeting peptide.
XX
XX Target; delivery; radionuclide chelator; diagnosis; therapy; detection;
KW atherosclerosis; thrombosis; Alzheimer's disease.
XX
OS Synthetic.
XX
XX US5659041-A.
XX
PD 19-AUG-1997.
XX
PF 02-SEP-1994; 94US-00299636.
XX
XX 19-JUL-1993; 93US-00092911.
XX
XX (RESO-) RESOLUTION PHARM INC.
PA
XX

PI Pollak A, Kirby RA, Dunn-Dufault R;
 XX WPI; 1997-424290/39.
 XX
 XX New thioacetamyl-aminocid hydrazide compounds - useful as chemical
 PT chelator of radionuclides for radio-imaging of target tissues of
 PT diagnostic interest.
 XX
 XX Disclosure; Col 29; 20pp; English.
 XX
 XX AAW31110-W31147 are peptides used for targeting a new hydrazone-type
 CC compound to various sites of disease, e.g. atherosclerotic plaque, sites
 CC of infection, platelets, thrombus or amyloid plaque. The new compound is
 CC a radionuclide chelator and is used to radiolabel the targeting peptides
 CC for the detection and diagnostic imaging of sites of disease, e.g.
 CC amyloid plaques in Alzheimer's disease. (Updated on 25-MAR-2003 to
 CC correct PF field.)
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 100.0%; Score 29; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 Db
 RESULT 3
 ID AAY49845 standard; peptide; 5 AA.
 AC AAY49845;
 XX
 DT 20-JAN-2000 (first entry)
 XX
 DE Tuftsin receptor antagonist peptide.
 XX
 KW Tuftsin receptor antagonist; chelate conjugate; radiopharmaceutical;
 KW diagnosis; infection; inflammation; imaging; cancer; tumour.
 XX
 OS Synthetic.
 XX
 PN WO9951628-A1.
 XX
 PD 14-OCT-1999.
 XX
 PF 29-MAR-1999; 99WO-US006824.
 XX
 PR 03-APR-1998; 98US-0080672P.
 XX
 XX (DUPO) DU PONT PHARM CO.
 XX
 XX Edwards DS, Rajopadhye M;
 XX
 XX WPI; 1999-633729/54.
 DR
 XX
 XX New polypeptides and radiopharmaceuticals used for imaging infection,
 PT inflammation and cancer.
 PT
 XX
 XX Disclosure; Page 3; 80pp; English.
 XX
 XX The present invention describes polypeptide compounds of formula (I),
 CC capable of direct transformation into a radiopharmaceutical: Ch-Ln-(X1-X2
 CC -X3-X4-X5)d (I), where X1-X5 = amino acids; Ln = a linking group; Ch = a
 CC metal bonding unit; and d is selected from 1, 2 and 3. The
 CC radiopharmaceuticals are useful for the diagnosis of infection,
 CC inflammation and cancer. The radiopharmaceuticals bind in vivo to the
 CC tuftsin receptor on the surface of white cells which accumulate at the
 CC site of infection and inflammation and can then be detected using
 CC radiation detecting probes or by imaging using a planar or ring gamma
 CC camera. The radiopharmaceuticals can also be used in treating cancer. The

CC present sequence represents the tuftsin receptor antagonist peptide given
 CC in the present invention
 XX
 SQ Sequence 5 AA;
 Query Match 100.0%; Score 29; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 Db
 RESULT 4
 ID ABB08442 standard; peptide; 5 AA.
 XX
 XX ABB08442;
 AC
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Tuftsin receptor antagonist (TKPPR).
 XX
 KW Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
 KW angiogenesis; tumour cell; cytostatic; antagonist.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "residue optionally modified by one of the
 FT following; Thr(Obzl), DPPE-Glutaryl-
 FT Di(aminodioxoactanoyl), aminodioxoactanoyl-Thr (Obzl),
 FT di(aminodioxoactanoyl)-Thr(Obzl), (t-Bu), Fmoc-Thr(t-Bu),
 FT H-Thr(t-Bu), Boc-Thr(But)"
 FT Modified-site 2 /note= "residue optionally modified by one of the
 FT following; Z (not further defined) or (Boc)"
 FT Modified-site 5 /note= "residue optionally modified by one of the
 FT following; OH, (NO2)Obzl, (Fmc)-t-Bu, (Fmc)-COHN"
 XX
 XX WO200191805-A2.
 XX
 XX 06-DEC-2001.
 XX
 XX 04-JUN-2001; 2001WO-US018053.
 XX
 XX 02-JUN-2000; 2000US-00585364.
 XX
 XX (BRAC) BRACCO RES USA.
 XX
 XX Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
 XX Tweedle MF, Linder K, Nanjappan P, Raju N;
 XX
 XX WPI; 2002-195523/25.
 DR
 XX
 XX Composition for use in targeting endothelial cells, tumour cells or other
 PT cells which express NP-1 comprises a compound containing a polypeptide,
 PT linker and substrate.
 PT
 XX
 XX Claim 1; Page 2; 146pp; English.
 PS
 XX
 XX The invention relates to a composition for use in targeting endothelial
 CC cells, tumour cells, or other cells which express NP-1. The activity of
 CC compositions of the invention may be described as cytostatic. Compounds
 CC of the invention are useful in pharmaceutical compositions for inhibiting
 CC angiogenesis, for imaging and targeting an angiogenic site, endothelial
 CC cells, tumour cells or other cells that express NP-1 in a human or
 CC animal. They may also be used as ultrasound contrast agents, for staging
 CC a tumour in a human or animal, for screening for the ability of an agent
 CC to target endothelial cells, tumour cells or other cells that express NP-

CC 1. They may be used for therapeutic delivery in vivo of a bioactive agent
 CC or for treating an individual exhibiting effects of an angiogenesis or a
 CC related disorder. They may be used for delivering desired nucleic acids
 CC to endothelial cells, tumour cells or other cells expressing NP-1, for
 CC enhancing endothelial or tumour cell-targeted gene therapy, or gene
 CC therapy targeting angiogenic cells, and for treating a human or animal
 CC with a tumour or angiogenesis-related disease. The current sequence
 CC represents the tuftsin receptor antagonist (TKPPR), of which monomers,
 CC multimers, polymers or analogues of, may be used to target endothelial
 CC cells, or cells that express markers in common with endothelial cells
 XX
 SQ Sequence 5 AA;

Query Match 100.0%; Score 29; DB 5; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06; Indels 0;
 Matches 5; Conservative 0; Mismatches 0; Gaps 0;

OY 1 TKPPR 5
 |||||
 DB 1 TKPPR 5

RESULT 5
 AAM51906
 ID AAM51906 standard; peptide; 5 AA.

AC AAM51906;
 DT 01-FEB-2002 (first entry)
 XX Tertiary ligand complex peptide #8.
 DE Ternary ligand complex; highly functionalised phosphine ligand;
 KW disease site; cardiovascular disorder; thromboembolic disease;
 KW atherosclerosis; infectious disease; cancer; radiopharmaceutical.
 XX
 OS Synthetic.

XX Key Location/Qualifiers
 FH Modified-site 1
 FT /label= OTHER
 FT /note= "modified by 6-aminohexanamide"

XX WO200177122-A1.

XX 18-OCT-2001.

XX 06-APR-2001; 2001WO-US011387.

XX 07-APR-2000; 2000US-0195235P.

XX (DUPO) DUPONT PHARM CO.

XX Liu S;

XX WPI; 2002-010884/01.

XX New ancillary ligands are useful as ligands for radiopharmaceuticals,
 PT which are useful for e.g. radioimaging a patient, diagnosing
 PT thromboembolic disorders, atherosclerosis, infections and inflammation.

XX Example; Page 126; 210pp; English.

XX The present invention relates to novel highly functionalised phosphine
 CC ligands which can be used as ancillary ligands in radiopharmaceuticals.
 CC These radiopharmaceuticals can then be used to radioimage a patient,
 CC enabling the visualisation of platelet deposition and disease sites, and
 CC allowing the diagnosis of infection, inflammation, transplant rejection,
 CC cardiovascular diseases such as thromboembolic disorders, and cancer.
 CC The present sequence is a peptide which formed part of a ligand in the
 CC exemplification of the invention

XX Sequence 5 AA;

Query Match 100.0%; Score 29; DB 5; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06; Indels 0;
 Matches 5; Conservative 0; Mismatches 0; Gaps 0;

OY 1 TKPPR 5
 |||||
 DB 1 TKPPR 5

RESULT 6
 ADD10684
 ID ADD10684 standard; peptide; 5 AA.

XX ADD10684;

XX 01-JAN-2004 (first entry)

XX Tuftsin analogue peptide #1.

XX Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
 KW antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
 KW visualisation therapy; radiotherapy.

XX Synthetic.

XX Key Location/Qualifiers
 FH Modified-site 1
 FT /label= OTHER
 FT /note= "Thr is optionally linked to an oregon green (OG)
 FT moiety"

XX US2002147136-A1.

XX 10-OCT-2002.

XX 04-JUN-2001; 2001US-00871974.

XX 02-JUN-2000; 2000US-00585364.

XX (VWRO// VON WRONSKI M A.

XX (MARI// MARINELLI E R.

XX (NUNN// NUNN A D.

XX (PILL// PILLAI R.

XX (RAMA// RAMALINGAM K.

XX (TWEE// TWEEDLE M F.

XX (LIND// LINDER K.

XX (NANJ// NANJAPPAN P.

XX (RAJU// RAJU N.

XX Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;

XX Tweedle MF, Linder K, Nanjappan P, Raju N;

XX WPI; 2003-800817/75.

XX Composition used in targeting endothelial cells e.g. tumor cells
 PT comprises compounds containing monomers, multimers or polymers of L-
 PT arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.

XX Claim 1; SEQ ID NO 2; 85pp; English.

XX The invention relates to a composition (A1) comprising compounds
 CC containing monomers, multimers or polymers of TKPPR (ADD10684).
 CC Composition (A1) comprises a compound of formula A-L-B 1, where A is the
 CC TKPPR peptide, L is a linker moiety (of formula given in the
 CC specification) and B is a substrate (or a phospholipid group,
 CC derivatisable bead attached to a fluorescent or radioactive marker,
 CC bioactive agent, delivery vehicle for genetic material, drug or
 CC therapeutic, or chelating group (preferably N 4, S 4, N 3 S, N 2 S 2 or
 CC NS 3) comprising oxo-pheno complexed with 9m rc). The compound
 CC specifically binds to NP-1 (Vascular endothelial growth factor binding
 CC receptor transmembrane glycoprotein) or cells that express NP-1 with
 CC avidity of at least that of TKPPR. Also included are an ultrasound

CC contrast agent (c1) comprising a suspension of gas filled microbubbles
 CC comprising the TKPPR compound, an ultrasound contrast agent (c2)
 CC comprising a suspension of gas filled microballoons comprising the TKPPR
 CC compound, preparation of the TKPPR compound (which comprises conjugating
 CC the monomer, multimer or polymer of TKPPR or its analogue with a linker
 CC to obtain a compound of formula A-L, forming a covalent or non-covalent
 CC bond between A-L and the substrate B-L or forming a covalent bond between
 CC B-L and the linker to form a conjugate B-L followed by conjugation with
 CC the monomer), and a kit for preparing a radiopharmaceutical comprising
 CC the compound. The compound used for targeting endothelial cells, tumour
 CC cells or other cells which express NP-1, for inhibiting angiogenesis, for
 CC ultrasound imaging, staging a tumour, screening at least one targeted
 CC ultrasound contrast agent for the ability to target endothelial cells,
 CC tumour cells or other cells which express NP-1, for the therapeutic
 CC delivery in vivo of a bioactive agent and for delivering desired nucleic
 CC acids to endothelial cells, tumour cells or other cells which express NP-
 CC 1. The composition is also useful for visualisation therapy or
 CC radiotherapy of endothelial cells. The present sequence is the TKPPR
 CC peptide.
 CC
 CC SQ Sequence 5 AA;

Query Match 100.0%; Score 29; DB 7; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TKPPR 5
 |||||
 DB 1 TKPPR 5

RESULT 7
 AAR85540
 ID AAR85540 standard; peptide; 6 AA.
 AC AAR85540;
 XX
 DT 17-APR-1996 (first entry)
 DE
 XX
 XX Tuftsin analogue conjugated to metal-chelating moiety.
 KW tuftsin; chelator; radiodiagnostic; imaging; leukocytes; radionuclide.
 OS Synthetic;
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "the amino group of this residue is acylated with
 FT (S-Acm-mercaptopoacetyl)-Ser- (N-methyl)hydrazino-
 FT nicotinoyl"
 FT
 FN WO9503280-A1.
 XX
 PD 02-FEB-1995.
 XX
 PF 18-JUL-1994; 94WO-CA000395.
 XX
 PR 19-JUL-1993; 93US-0002911.
 XX
 PA (RESO-) RESOLUTION PHARM INC.
 XX
 XX Pollak A, Kirby RA, Dufault R;
 XX WPI; 1995-075166/10.
 DR
 XX

XX Hydrazino-type radionuclide chelating agents - capable of conjugation to
 PT targeting molecules such as proteins, peptide(s) or antibodies and thus
 PT useful in diagnosis and therapy.
 XX
 XX Claim 29; Page 29; 35pp; English.

CC The patent provides radionuclide chelating compounds and their
 CC conjugation products with targeting molecules such as proteins, peptides

CC or antibodies. The labelled targeting molecules may be used for in-vivo
 CC diagnosis and therapy. The chelating compound is a hydrazino-containing,
 CC N-heterocycle-containing compound, and the conjugated product is
 CC exemplified by S-Acm-mercaptopoacetyl-Ser-N-methyl-hydrazino-nicotinic
 CC acid-Gly-Thr-Lys-Pro-Arg. The peptide portion of this molecule (the
 CC present sequence) represents the tuftsin antagonist TKPPR condensed via
 CC Gly onto the chelating moiety, giving a molecule which will chelate a
 CC radionuclide such as 99mTc and target it towards leukocytes
 XX
 CC SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TKPPR 5
 |||||
 DB 2 TKPPR 6

RESULT 8
 AAW11053
 ID AAW11053 standard; peptide; 6 AA.
 AC AAW11053;
 XX
 DT 03-JUN-1997 (first entry)
 DE
 XX Leukocyte-targeting peptide used in diagnostic imaging.
 DE Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
 KW diagnostic imaging.
 XX
 XX Synthetic.
 XX
 XX WO9603427-A1.
 XX
 PD 08-FEB-1996.
 XX
 PF 28-APR-1995; 95WO-CA000249.
 XX
 PR 22-JUL-1994; 94US-00279155.
 XX
 PA (RESO-) RESOLUTION PHARM INC.
 XX
 XX Pollak A, Goodbody A;
 PI
 DR WPI; 1996-116994/12.
 XX
 PT New peptide derived radionuclide chelators and metal complexes - useful
 PT for diagnostic imaging.
 XX
 PS Claim 13; Page 20; 30pp; English.

XX AAW11053 is a peptide used for targeting agents to leukocytes. This
 CC peptide can be coupled to a metal radionuclide chelator (structural
 CC formula given in the specification) labelled with a diagnostically
 CC useful metal isotope, to form a peptide derived radionuclide chelator.
 CC When the chelator is coupled to a targeting molecule and labelled with a
 CC diagnostically useful metal, it can be used to detect pathological
 CC conditions by diagnostic imaging. For example, AAW11053 targets the
 CC chelator to leukocytes and are useful for the rapid imaging of sites of
 CC local inflammation. Radionuclides used include 99mTc, 64Cu, 67Cu, 97Ru,
 CC 105Rh, 109Pd, 186Re, 188Re, 198Au, 203Pb, 212Pb and 212Bi. The
 CC coupling of a targeting agent and radionuclide using a chelating agent is
 CC an alternative to the direct labelling of targeting agents in which
 CC radionuclides are typically bound at the more numerous low-affinity
 CC sites, forming unstable complexes. The new conjugates give better
 CC scintigraphic images in rat inflammation studies than known imaging
 CC agents Ga-67, 99mTc-IGG, 111In-WBC and 99mTc-Nanocol. They image more
 CC rapidly than the known agents and show superior biodistribution
 XX
 CC SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TKPPR 5
 Db 2 TKPPR 6

RESULT 9
 AAW31148
 ID AAW31148 standard; peptide; 6 AA.
 XX AC AAW31148;
 XX DT 25-MAR-2003 (revised)
 XX DT 23-JAN-1998 (first entry)
 XX DE Amyloid plaque-targeting peptide-radiolabeled chelator conjugate.
 XX DE Target; delivery; radiolabeled chelator; diagnosis; therapy; detection;
 XX KW atherosclerosis; thrombosis; Alzheimer's disease.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "N terminally acetylated with S-Acm-Mercaptoacetyl
 FT -Ser-N- methylhydrazino nicotinic acid"
 XX US5659041-A.
 XX PD 19-AUG-1997.
 XX PF 02-SEP-1994; 94US-00299636.
 XX PR 19-JUL-1993; 93US-00092911.
 XX PA (RESO-) RESOLUTION PHARM INC.
 XX PI Pollak A, Kirby RA, Dunn-Dufault R;
 XX DR WPI; 1997-424290/39.
 XX PT New thio:acetyl-aminoacid hydrazide compounds - useful as chemical
 XX PT chelator of radionuclides for radio:imaging of target tissues of
 XX PT diagnostic interest.
 XX PS Example 11; Col 14; 20pp; English.
 XX CC AAW31148 is a peptide conjugate of an amyloid plaque-targeting peptide
 XX CC and a new hydrazino-type compound. The new compound is a radionuclide
 XX CC chelator and is used to radiolabel targeting peptides for the detection
 XX CC and diagnostic imaging of sites of disease, e.g. amyloid plaques in
 XX CC Alzheimer's disease, thrombi in thrombosis or other sites of infection.
 XX CC (Updated on 25-MAR-2003 to correct PF field.)
 XX SQ Sequence 6 AA;
 Query Match 100.0%; Score 29; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TKPPR 5
 Db 2 TKPPR 6

RESULT 10
 AAW31147
 ID AAW31147 standard; peptide; 6 AA.
 XX AC AAW31147;
 XX DT 25-MAR-2003 (revised)
 XX DT 23-JAN-1998 (first entry)
 XX DE Amyloid plaque-targeting peptide.
 XX DE Target; delivery; radionuclide chelator; diagnosis; therapy; detection;
 XX KW atherosclerosis; thrombosis; Alzheimer's disease.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1
 FT /label= Acp
 FT /note= "modified 6-aminocaproic acid: 6-((6-((1-aza-2-
 FT -sulphophenyl)vinyl)amino)-3-pyridyl)
 FT carbonylamino)hexanoyl"
 XX US5659041-A.
 XX PD 19-AUG-1997.
 XX PF 02-SEP-1994; 94US-00299636.
 XX PR 19-JUL-1993; 93US-00092911.
 XX PA (RESO-) RESOLUTION PHARM INC.
 XX PI Pollak A, Kirby RA, Dunn-Dufault R;
 XX DR WPI; 1997-424290/39.
 XX PT New thio:acetyl-aminoacid hydrazide compounds - useful as chemical
 XX PT chelator of radionuclides for radio:imaging of target tissues of
 XX PT diagnostic interest.
 XX PS Disclosure; Col 29; 20pp; English.
 XX CC AAW3110-W31147 are peptides used for targeting a new hydrazino-type
 XX CC compound to various sites of disease, e.g. atherosclerotic plaque, sites
 XX CC of infection, platelets, thrombus or amyloid plaque. The new compound is
 XX CC a radionuclide chelator and is used to radiolabel the targeting peptides
 XX CC for the detection and diagnostic imaging of sites of disease, e.g.
 XX CC amyloid plaques in Alzheimer's disease. (Updated on 25-MAR-2003 to
 XX CC correct PF field.)
 XX SQ Sequence 6 AA;
 Query Match 100.0%; Score 29; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TKPPR 5
 Db 2 TKPPR 6

RESULT 11
 AAY49840
 ID AAY49840 standard; peptide; 6 AA.
 XX AC AAY49840;
 XX DT 20-JAN-2000 (first entry)
 XX DE Tuftsin receptor antagonist chelate conjugate #1.
 XX KW Tuftsin receptor antagonist; chelate conjugate; radiopharmaceutical;
 XX KW diagnosis; infection; inflammation; imaging; cancer; tumour.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1
 FT /label= Acp
 FT /note= "modified 6-aminocaproic acid: 6-((6-((1-aza-2-
 FT -sulphophenyl)vinyl)amino)-3-pyridyl)
 FT carbonylamino)hexanoyl"
 XX US5659041-A.
 XX PD 19-AUG-1997.
 XX PF 02-SEP-1994; 94US-00299636.
 XX PR 19-JUL-1993; 93US-00092911.
 XX PA (RESO-) RESOLUTION PHARM INC.
 XX PI Pollak A, Kirby RA, Dunn-Dufault R;
 XX DR WPI; 1997-424290/39.
 XX PT New thio:acetyl-aminoacid hydrazide compounds - useful as chemical
 XX PT chelator of radionuclides for radio:imaging of target tissues of
 XX PT diagnostic interest.
 XX PS Example 11; Col 14; 20pp; English.
 XX CC AAW31148 is a peptide conjugate of an amyloid plaque-targeting peptide
 XX CC and a new hydrazino-type compound. The new compound is a radionuclide
 XX CC chelator and is used to radiolabel targeting peptides for the detection
 XX CC and diagnostic imaging of sites of disease, e.g. amyloid plaques in
 XX CC Alzheimer's disease, thrombi in thrombosis or other sites of infection.
 XX CC (Updated on 25-MAR-2003 to correct PF field.)
 XX SQ Sequence 6 AA;
 Query Match 100.0%; Score 29; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TKPPR 5
 Db 2 TKPPR 6

RESULT 11
 AAW31147
 ID AAW31147 standard; peptide; 6 AA.
 XX AC AAW31147;
 XX DT 25-MAR-2003 (revised)
 XX DT 23-JAN-1998 (first entry)
 XX DE Amyloid plaque-targeting peptide.
 XX DE Target; delivery; radionuclide chelator; diagnosis; therapy; detection;
 XX KW atherosclerosis; thrombosis; Alzheimer's disease.
 XX OS Synthetic.
 XX FH Key Location/Qualifiers
 FT Modified-site 1
 FT /label= Acp
 FT /note= "modified 6-aminocaproic acid: 6-((6-((1-aza-2-
 FT -sulphophenyl)vinyl)amino)-3-pyridyl)
 FT carbonylamino)hexanoyl"
 XX US5659041-A.
 XX PD 19-AUG-1997.
 XX PF 02-SEP-1994; 94US-00299636.
 XX PR 19-JUL-1993; 93US-00092911.
 XX PA (RESO-) RESOLUTION PHARM INC.
 XX PI Pollak A, Kirby RA, Dunn-Dufault R;
 XX DR WPI; 1997-424290/39.
 XX PT New thio:acetyl-aminoacid hydrazide compounds - useful as chemical
 XX PT chelator of radionuclides for radio:imaging of target tissues of
 XX PT diagnostic interest.
 XX PS Disclosure; Col 29; 20pp; English.
 XX CC AAW3110-W31147 are peptides used for targeting a new hydrazino-type
 XX CC compound to various sites of disease, e.g. atherosclerotic plaque, sites
 XX CC of infection, platelets, thrombus or amyloid plaque. The new compound is
 XX CC a radionuclide chelator and is used to radiolabel the targeting peptides
 XX CC for the detection and diagnostic imaging of sites of disease, e.g.
 XX CC amyloid plaques in Alzheimer's disease. (Updated on 25-MAR-2003 to
 XX CC correct PF field.)
 XX SQ Sequence 6 AA;
 Query Match 100.0%; Score 29; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TKPPR 5
 Db 2 TKPPR 6

PN WO9951628-A1.
 XX
 PD 14-OCT-1999.
 XX
 PF 29-MAR-1999; 99WO-US006824.
 XX
 PR 03-APR-1998; 98US-0080672P.
 XX
 PA (DUPO) DU PONT PHARM CO.
 XX
 XX Edwards DS, Rajopadhye M;
 PI WPI; 1999-633729/54.
 XX
 DR New polypeptides and radiopharmaceuticals used for imaging infection,
 XX inflammation and cancer.
 PT
 PT
 XX
 PS Claim 4; Page 71; 80pp; English.
 XX
 CC The present invention describes polypeptide compounds of formula (I),
 CC capable of direct transformation into a radiopharmaceutical: $\text{Ch-Ln-(X1-X2-X3-X4-X5)d}$ (I), where X1-X5 = amino acids; Ln = a linking group; Ch = a
 CC metal bonding unit; and d is selected from 1, 2 and 3. The
 CC radiopharmaceuticals are useful for the diagnosis of infection,
 CC inflammation and cancer. The radiopharmaceuticals bind in vivo to the
 CC tuftsin receptor on the surface of white cells which accumulate at the
 CC site of infection and inflammation and can then be detected using
 CC radiation detecting probes or by imaging using a planar or ring gamma
 CC camera. The radiopharmaceuticals can also be used in treating cancer. The
 CC present sequence represents a specifically claimed tuftsin receptor
 CC antagonist chelate conjugate from the present invention
 CC
 XX
 SQ Sequence 6 AA;
 Query Match 100.0%; Score 29; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 2 TKPPR 6
 RESULT 12
 ABB08444
 ID ABB08444 standard; peptide; 6 AA.
 AC ABB08444;
 XX
 XX 01-JUL-2002 (first entry)
 DE Tuftsin receptor antagonist (TKPPR) derivative peptide 2.
 XX
 KW Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
 KW angiogenesis; tumour cell; cytostatic; antagonist.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "residue optionally modified by one of the
 FT following; DPPG-Glutaryl, N-Glutaryl-Gly, NH2, Oregon
 FT Modified-site 2
 FT /note= "residue optionally modified by one of the
 FT following; Obz1, (tBu)"
 FT Modified-site 3
 FT /note= "residue optionally modified by one of the
 FT following; 2 (not further defined), (Mtt), (Boc)"
 FT Modified-site 6
 FT /note= "residue optionally modified by one of the
 FT following; OH, (NO2)Obz1, (NO2), (Pmc)Obz1, (pmc)-wang-
 FT resin"

XX
 PN WO200191805-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 04-JUN-2001; 2001WO-US018053.
 XX
 PR 02-JUN-2000; 2000US-00585364.
 XX
 PA (BRAC) BRACCO RES USA.
 XX
 XX Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
 PI Tweedle MF, Linder K, Narjappa P, Raju N;
 XX WPI; 2002-195523/25.
 DR
 XX
 PT Composition for use in targeting endothelial cells, tumor cells or other
 PT cells which express NP-1 comprises a compound containing a polypeptide,
 PT linker and substrate.
 XX
 XX Example 4; Page 67; 145pp; English.
 XX
 CC The invention relates to a composition for use in targeting endothelial
 CC cells, tumour cells, or other cells which express NP-1. The activity of
 CC compositions of the invention may be described as cytostatic. Compounds
 CC of the invention are useful in pharmaceutical compositions for inhibiting
 CC angiogenesis, for imaging and targeting an angiogenic site, endothelial
 CC cells, tumour cells or other cells that express NP-1 in a human or
 CC animal. They may also be used as ultrasound contrast agents, for staging
 CC a tumour in a human or animal, for screening for the ability of an agent
 CC to target endothelial cells, tumour cells or other cells that express NP-1.
 CC 1. They may be used for therapeutic delivery in vivo of a bioactive agent
 CC or for treating an individual exhibiting effects of an angiogenesis or a
 CC related disorder. They may be used for delivering desired nucleic acids
 CC to endothelial cells, tumour cells or other cells expressing NP-1, for
 CC enhancing endothelial or tumour cell-targeted gene therapy, or gene
 CC therapy targeting angiogenic cells, and for treating a human or animal
 CC with a tumour or angiogenesis-related disease. The current sequence
 CC represents a tuftsin receptor antagonist (TKPPR) derivative of the
 CC invention
 XX
 SQ Sequence 6 AA;
 Query Match 100.0%; Score 29; DB 5; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 2 TKPPR 6
 RESULT 13
 AAM51904
 ID AAM51904 standard; peptide; 6 AA.
 AC AAM51904;
 XX
 XX 01-FEB-2002 (first entry)
 DE
 DT
 XX
 XX Tertiary ligand complex peptide #6.
 DE
 XX Ternary ligand complex; highly functionalised phosphine ligand;
 KW disease site; cardiovascular disorder; thromboembolic disease;
 KW atherosclerosis; infectious disease; cancer; radiopharmaceutical.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 1
 FT /note= "D-form residue"
 FT Modified-site 2
 FT /label= OTHER

/note= "modified by 6-aminohexanamide"

PA (DUPO) DUPONT PHARM CO.
 XX
 PI Liu S;
 XX
 DR WPI; 2002-010884/01.
 XX
 PT New ancillary ligands are useful as ligands for radiopharmaceuticals,
 which are useful for e.g. radioimaging a patient, diagnosing
 PT thromboembolic disorders, atherosclerosis, infections and inflammation.
 XX
 PS Example; Page 125; 210pp; English.
 XX
 CC The present invention relates to novel highly functionalised phosphine
 CC ligands which can be used as ancillary ligands in radiopharmaceuticals.
 CC These radiopharmaceuticals can then be used to radioimage a patient,
 CC enabling the visualisation of platelet deposition and disease sites, and
 CC allowing the diagnosis of infection, inflammation, transplant rejection,
 CC cardiovascular diseases such as thromboembolic disorders, and cancer. The
 CC present sequence is a peptide which formed part of a ligand in the
 CC exemplification of the invention
 XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 5; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB |||||
 2 TKPPR 6

RESULT 15

ADD10689
 ID ADD10689 standard; peptide; 6 AA.

XX
 AC ADD10689;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Tuftsin analogue peptide #4.
 XX
 KW Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
 KW antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
 KW visualisation therapy; radiotherapy.
 XX
 OS Synthetic.
 XX
 PH Key Location/Qualifiers
 FT Modified-site 1 /label= OTHER
 FT /note= "Glu is covalently linked to a DPPE
 (dibenzoylphosphatidylethanolamine) moiety"
 FT Modified-site 2 /label= OTHER
 FT /note= "Thr is di(aminodioxanoyl)-Thr"

US2002147136-A1.

10-OCT-2002.

04-JUN-2001; 2001US-00871974.

02-JUN-2000; 2000US-00585364.

(VWRO/) VON WRONSKI M A.

(MARI/) MARINELLI E R.

(NUNN/) NUNN A D.

(PILL/) PILLAI R.

(RAMA/) RAMALINGAM K.

(TWEED/) TWEEDLE M F.

(LIND/) LINDER K.

(NANJ/) NANJAPPAN P.

FT
 XX
 PN WO200177122-A1.
 XX
 PD 18-OCT-2001.

XX
 PF 06-APR-2001; 2001WO-US011387.
 XX
 PR 07-APR-2000; 2000US-0195235P.

XX
 PA (DUPO) DUPONT PHARM CO.
 XX
 PI Liu S;
 XX
 DR WPI; 2002-010884/01.

XX
 PT New ancillary ligands are useful as ligands for radiopharmaceuticals,
 PT which are useful for e.g. radioimaging a patient, diagnosing
 PT thromboembolic disorders, atherosclerosis, infections and inflammation.
 XX
 PS Example; Page 125; 210pp; English.

XX
 CC The present invention relates to novel highly functionalised phosphine
 CC ligands which can be used as ancillary ligands in radiopharmaceuticals.
 CC These radiopharmaceuticals can then be used to radioimage a patient,
 CC enabling the visualisation of platelet deposition and disease sites, and
 CC allowing the diagnosis of infection, inflammation, transplant rejection,
 CC cardiovascular diseases such as thromboembolic disorders, and cancer. The
 CC present sequence is a peptide which formed part of a ligand in the
 CC exemplification of the invention
 XX
 SQ Sequence 6 AA;

Query Match 100.0%; Score 29; DB 5; Length 6;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB |||||
 2 TKPPR 6

RESULT 14

AAM51905
 ID AAM51905 standard; peptide; 6 AA.

XX
 AC AAM51905;
 XX
 DT 01-FEB-2002 (first entry)
 XX
 DE Tertiary ligand complex peptide #7.
 XX
 KW Ternary ligand complex; highly functionalised phosphine ligand;
 KW disease site; cardiovascular disorder; thromboembolic disease;
 KW atherosclerosis; infectious disease; cancer; radiopharmaceutical.
 XX
 OS Synthetic.

PH Key Location/Qualifiers
 FT Misc-difference 1 /note= "D-form residue"

FT Modified-site 2 /label= OTHER

FT /note= "modified by 6-aminohexanamide"

WO200177122-A1.

18-OCT-2001.

06-APR-2001; 2001WO-US011387.

07-APR-2000; 2000US-0195235P.

PA (RAJU//) RAJU N.
 XX Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
 PI Tweedle MF, Linder K, Nanjappan P, Raju N;
 XX WPI; 2003-800817/75.
 XX Composition used in targeting endothelial cells e.g. tumor cells
 PT comprises compounds containing monomers, multimers or polymers of L-
 PT arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.
 XX Example 23; Page 44; 85pp; English.
 XX The invention relates to a composition (A1) comprising compounds
 CC containing monomers, multimers or polymers of TKPPR (ADD10684).
 CC Composition (A1) comprises a compound of formula A-L-B 1, where A is the
 CC TKPPR peptide, L is a linker moiety (of formula given in the
 CC specification), and B is a substrate (or a phospholipid group,
 CC derivatisable bead attached to a fluorescent or radioactive marker,
 CC bioactive agent, delivery vehicle for genetic material, drug or
 CC therapeutic, or chelating group (preferably N 4, S 4, N 3 S, N 2 S 2 or
 CC NS 3) comprising oxa-PnAO complexed with 9m Tc). The compound
 CC specifically binds to NP-1 (Vascular endothelial growth factor binding
 CC receptor transmembrane glycoprotein) or cells that express NP-1 with
 CC avidity of at least that of TKPPR. Also included are an ultrasound
 CC contrast agent (c1) comprising a suspension of gas filled microbubbles
 CC comprising the TKPPR compound, an ultrasound contrast agent (c2)
 CC comprising a suspension of gas filled microballoons comprising the TKPPR
 CC compound, preparation of the TKPPR compound (which comprises conjugating
 CC the monomer, multimer or polymer of TKPPR or its analogue with a linker
 CC to obtain a compound of formula A-L, forming a covalent or non-covalent
 CC bond between A-L and the substrate B 1 or forming a covalent bond between
 CC B 1 and the linker to form a conjugate B-L followed by conjugation with
 CC the monomer), and a kit for preparing a radiopharmaceutical comprising
 CC the compound. The compound used for targeting endothelial cells, tumour
 CC cells or other cells which express NP-1, for inhibiting angiogenesis, for
 CC ultrasound imaging, staging a tumour, screening at least one targeted
 CC ultrasound contrast agent for the ability to target endothelial cells,
 CC tumour cells or other cells which express NP-1, for the therapeutic
 CC delivery in vivo of a bioactive agent and for delivering desired nucleic
 CC acids to endothelial cells, tumour cells or other cells which express NP-
 CC 1. The composition is also useful for visualisation therapy or
 CC radiotherapy of endothelial cells. The present sequence is a TKPPR
 CC analogue peptide.
 XX Sequence 6 AA;
 Query Match 100.0%; Score 29; DB 7; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 2 TKPPR 6
 RESULT 16
 ADD10687
 ID ADD10687 standard; peptide; 6 AA.
 XX
 AC ADD10687;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Tuftsin analogue peptide #3.
 XX
 KW Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
 KW antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
 KW visualisation therapy; radiotherapy.
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers

FT Modified-site 1 /label= OTHER
 FT /note= "Gly is covalently linked to a DPPE
 FT (dipalmitoylphosphatidylethanolamine) moiety"
 XX
 PN US2002147136-A1.
 XX
 PD 10-OCT-2002.
 XX
 PF 04-JUN-2001; 2001US-00871974.
 XX
 PR 02-JUN-2000; 2000US-00585364.
 XX (VWRO//) VON WRONSKI M A.
 PA (MARI//) MARINELLI E R.
 PA (NUNN//) NUNN A D.
 PA (PILL//) PILLAI R.
 PA (RAWA//) RAMALINGAM K.
 PA (TWEE//) TWEEDLE M F.
 PA (LIND//) LINDER K.
 PA (NANJ//) NANJAPPAN P.
 PA (RAJU//) RAJU N.
 XX Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
 PI Tweedle MF, Linder K, Nanjappan P, Raju N;
 XX WPI; 2003-800817/75.
 CC Composition used in targeting endothelial cells e.g. tumor cells
 CC comprises compounds containing monomers, multimers or polymers of L-
 CC arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.
 CC Example 14; Page 38; 85pp; English.
 CC The invention relates to a composition (A1) comprising compounds
 CC containing monomers, multimers or polymers of TKPPR (ADD10684).
 CC Composition (A1) comprises a compound of formula A-L-B 1, where A is the
 CC TKPPR peptide, L is a linker moiety (of formula given in the
 CC specification), and B is a substrate (or a phospholipid group,
 CC derivatisable bead attached to a fluorescent or radioactive marker,
 CC bioactive agent, delivery vehicle for genetic material, drug or
 CC therapeutic, or chelating group (preferably N 4, S 4, N 3 S, N 2 S 2 or
 CC NS 3) comprising oxa-PnAO complexed with 9m Tc). The compound
 CC specifically binds to NP-1 (Vascular endothelial growth factor binding
 CC receptor transmembrane glycoprotein) or cells that express NP-1 with
 CC avidity of at least that of TKPPR. Also included are an ultrasound
 CC contrast agent (c1) comprising a suspension of gas filled microbubbles
 CC comprising the TKPPR compound, an ultrasound contrast agent (c2)
 CC comprising a suspension of gas filled microballoons comprising the TKPPR
 CC compound, preparation of the TKPPR compound (which comprises conjugating
 CC the monomer, multimer or polymer of TKPPR or its analogue with a linker
 CC to obtain a compound of formula A-L, forming a covalent or non-covalent
 CC bond between A-L and the substrate B 1 or forming a covalent bond between
 CC B 1 and the linker to form a conjugate B-L followed by conjugation with
 CC the monomer), and a kit for preparing a radiopharmaceutical comprising
 CC the compound. The compound used for targeting endothelial cells, tumour
 CC cells or other cells which express NP-1, for inhibiting angiogenesis, for
 CC ultrasound imaging, staging a tumour, screening at least one targeted
 CC ultrasound contrast agent for the ability to target endothelial cells,
 CC tumour cells or other cells which express NP-1, for the therapeutic
 CC delivery in vivo of a bioactive agent and for delivering desired nucleic
 CC acids to endothelial cells, tumour cells or other cells which express NP-
 CC 1. The composition is also useful for visualisation therapy or
 CC radiotherapy of endothelial cells. The present sequence is a TKPPR
 CC analogue peptide.
 XX Sequence 6 AA;
 Query Match 100.0%; Score 29; DB 7; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 2 TKPPR 6
 RESULT 16
 ADD10687
 ID ADD10687 standard; peptide; 6 AA.
 XX
 AC ADD10687;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Tuftsin analogue peptide #3.
 XX
 KW Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
 KW antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
 KW visualisation therapy; radiotherapy.
 XX Synthetic.
 OS
 XX
 FH Key Location/Qualifiers

Query Match 100.0%; Score 29; DB 7; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5

```
Db          |||||
            2 TKPPR 6

RESULT 17
AA49842
ID AAY49842 standard; peptide; 7 AA.
XX
XX AAY49842;
AC
XX
XX 20-JAN-2000 (first entry)
DT
XX
XX Tuftsin receptor antagonist chelate conjugate #3.
DE
XX
XX Tuftsin receptor antagonist; chelate conjugate; radiopharmaceutical;
KW diagnosis; infection; inflammation; imaging; cancer; tumour.
KW
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "modified tyrosine: ((6-((1-aza-2- (2-
FT sulphophenyl)vinyl)amino)-3-pyridyl) carbonyl)-L-
FT phenylalanine"
FT Modified-site 2
FT /label= Acp
FT
FT WO9951628-A1.
XX
XX 14-OCT-1999.
PN
XX
XX 29-MAR-1999; 99WO-US006824.
PF
XX
XX 03-APR-1998; 98US-0080672P.
PR
XX
XX (DUPO ) DU PONT PHARM CO.
PA
XX
XX Edwards DS, Rajopadhye M;
PI
XX
XX WPI; 1999-633729/54.
DR
XX
XX New polypeptides and radiopharmaceuticals used for imaging infection,
PT inflammation and cancer.
PS
XX
XX Claim 4; Page 71; 80pp; English.
XX
XX The present invention describes polypeptide compounds of formula (I),
XX capable of direct transformation into a radiopharmaceutical: Ch-Ln-(X1-X2
XX -X3-X4-X5)d (I), where X1-X5 = amino acids; Ln = a linking group; Ch = a
XX metal bonding unit; and d is selected from 1, 2 and 3. The
XX radiopharmaceuticals are useful for the diagnosis of infection,
XX inflammation and cancer. The radiopharmaceuticals bind in vivo to the
XX tuftsin receptor on the surface of white cells which accumulate at the
XX site of infection and inflammation and can then be detected using
XX radiation detecting probes or by imaging using a planar or ring gamma
XX camera. The radiopharmaceuticals can also be used in treating cancer. The
XX present sequence represents a specifically claimed tuftsin receptor
XX antagonist chelate conjugate from the present invention
SQ Sequence 7 AA;
Query Match 100.0%; Score 29; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
Db |||||
3 TKPPR 7

RESULT 18
AA49841
ID AAY49841 standard; peptide; 7 AA.
XX
XX Tuftsin receptor antagonist (TKPPR) derivative dimer.
DE
XX
```

KW Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
 KW angiogenesis; tumour cell; cytostatic; antagonist.
 XX Synthetic.

XX Key Location/Qualifiers
 XX Modified-site 6
 FT /note= "residue modified by the addition of the following
 FT block of residues (RPPKT)"
 FT Modified-site 7
 FT /note= "residue modified by TTD-A-OG (TTDA= 4,7,10-Trioxa-
 FT 1,13-TridecaneDiamine, OG= Oregon green)"

XX WO200191805-A2.

XX 06-DEC-2001.

XX 04-JUN-2001; 2001WO-US018053.

XX 02-JUN-2000; 2000US-00585364.

XX (BRAC) BRACCO RES USA.

XX Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
 PI Tweedle MF, Linder K, Nanjappan P, Raju N;

XX WPI; 2002-195523/25.

XX Composition for use in targeting endothelial cells, tumor cells or other
 PT cells which express NP-1 comprises a compound containing a polypeptide,
 PT linker and substrate.

XX Example 29; Page 113; 146pp; English.

XX The invention relates to a composition for use in targeting endothelial
 CC cells, tumour cells, or other cells which express NP-1. The activity of
 CC compositions of the invention may be described as cytostatic. Compounds
 CC of the invention are useful in pharmaceutical compositions for inhibiting
 CC angiogenesis, for imaging and targeting an angiogenic site, endothelial
 CC cells, tumour cells or other cells that express NP-1 in a human or
 CC animal. They may also be used as ultrasound contrast agents, for staging
 CC a tumour in a human or animal, for screening for the ability of an agent
 CC to target endothelial cells, tumour cells or other cells that express NP-
 CC 1. They may be used for therapeutic delivery in vivo of a bioactive agent
 CC or for creating an individual exhibiting effects of an angiogenesis or a
 CC related disorder. They may be used for delivering desired nucleic acids
 CC to endothelial cells, tumour cells or other cells expressing NP-1, for
 CC enhancing endothelial or tumour cell-targeted gene therapy, or gene
 CC therapy targeting angiogenic cells, and for treating a human or animal
 CC with a tumour or angiogenesis-related disease. The current sequence
 CC represents a tuftsin receptor antagonist (TKPPR) derivative dimer of the
 CC invention

XX Sequence 7 AA;

Query Match 100.0%; Score 29; DB 5; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

DB 1 TKPPR 5

RESULT 20

AB08445

ID AB08445 standard; peptide; 7 AA.

XX AC AB08445;

XX 01-JUL-2002 (first entry)

XX Linear and cyclic CTKEPR (BRU-305 and BRU-306).

XX Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
 KW angiogenesis; tumour cell; cytostatic; antagonist.
 XX Synthetic.

XX Key Location/Qualifiers
 XX Disulfide_bond 1.7
 FT /note= "optionally present"

FT Modified-site 1
 FT /note= "Fmoc-Cys(Ttt)-OH, residue optionally modified by
 FT H group"

FT Modified-site 2

FT /note= "Fmoc-Thr(But)-OH"

FT Modified-site 3

FT /note= "Fmoc-Lys(Boc)-OH"

FT Modified-site 4

FT /note= "Fmoc-Pro-OH"

FT Modified-site 5

FT /note= "Fmoc-Pro-OH"

FT Modified-site 6

FT /note= "Fmoc-Arg(Pmc)-OH"

FT Modified-site 7

FT /note= "Fmoc-Cys(Ttt)-OH; residue optionally modified by
 FT an OH group"

XX WO200191805-A2.

XX 06-DEC-2001.

XX 04-JUN-2001; 2001WO-US018053.

XX 02-JUN-2000; 2000US-00585364.

XX (BRAC) BRACCO RES USA.

XX Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
 PI Tweedle MF, Linder K, Nanjappan P, Raju N;

XX WPI; 2002-195523/25.

XX Composition for use in targeting endothelial cells, tumor cells or other
 PT cells which express NP-1 comprises a compound containing a polypeptide,
 PT linker and substrate.

XX Example 21; Page 81; 146pp; English.

XX The invention relates to a composition for use in targeting endothelial
 CC cells, tumour cells, or other cells which express NP-1. The activity of
 CC compositions of the invention may be described as cytostatic. Compounds
 CC of the invention are useful in pharmaceutical compositions for inhibiting
 CC angiogenesis, for imaging and targeting an angiogenic site, endothelial
 CC cells, tumour cells or other cells that express NP-1 in a human or
 CC animal. They may also be used as ultrasound contrast agents, for staging
 CC a tumour in a human or animal, for screening for the ability of an agent
 CC to target endothelial cells, tumour cells or other cells that express NP-
 CC 1. They may be used for therapeutic delivery in vivo of a bioactive agent
 CC or for treating an individual exhibiting effects of an angiogenesis or a
 CC related disorder. They may be used for delivering desired nucleic acids
 CC to endothelial cells, tumour cells or other cells expressing NP-1, for
 CC enhancing endothelial or tumour cell-targeted gene therapy, or gene
 CC therapy targeting angiogenic cells, and for treating a human or animal
 CC with a tumour or angiogenesis-related disease. The current sequence
 CC represents a tuftsin receptor antagonist (TKPPR) derivative of the
 CC invention referred to as linear and cyclic CTKEPR (BRU-305 and BRU-306)

XX Sequence 7 AA;

Query Match 100.0%; Score 29; DB 5; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

Db 2 TKPPR 6

RESULT 21
ADD10688

ID ADD10688 standard; peptide; 7 AA.
AC ADD10688;
XX
DT 01-JAN-2004 (first entry)
XX
DE Cyclic Tuftsin analogue peptide.
XX
KW Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
KW antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
KW visualisation therapy; radiotherapy; cyclic.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FH Disulfide-bond 1..7
FT /note= "The molecule is cyclised via this disulfide bond"
FT
PN US2002147136-A1.
XX
XX 10-OCT-2002.
XX
XX 04-JUN-2001; 2001US-00871974.
XX
XX 02-JUN-2000; 2000US-00585364.
XX
XX (VWRO/) VON WRONSKI M A.
XX (WARI/) MARINELLI E R.
XX (NUNN/) NUNN A D.
XX (PILL/) PILLAI R.
XX (RAMA/) RAMALINGAM K.
XX (TWE/) TWEEDLE M F.
XX (LIND/) LINDER K.
XX (NANJ/) NANJAPPAN P.
XX (RAJU/) RAJU N.
XX
XX Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
XX Tweedle MF, Linder K, Nanjappan P, Raju N;
XX WPI; 2003-800817/75.
XX
XX Composition used in targeting endothelial cells e.g. tumor cells
XX comprises compounds containing monomers, multimers or polymers of L-
XX arginine-L-threonine-L-lysyl-L-prolyl-L-prolyl.
XX
XX Example 21; Page 41; 85pp; English.
XX
XX The invention relates to a composition (A1) comprising compounds
XX containing monomers, multimers or polymers of TKPPR (ADD10684).
XX Composition (A1) comprises a compound of formula A-L-B 1, where A is the
XX TKPPR peptide, L is a linker moiety (of formula given in the
XX specification) and B is a substrate (or a phospholipid group,
XX derivatisable bead attached to a fluorescent or radioactive marker,
XX bioactive agent, delivery vehicle for genetic material, drug or
XX therapeutic, or chelating group (preferably N 4, S 4, N 3 S, N 2 S 2 or
XX NS 3) comprising oxa-PnAO complexed with 99m Tc). The compound
XX specifically binds to NP-1 (Vascular endothelial growth factor binding
XX receptor transmembrane glycoprotein) or cells that express NP-1 with
XX avidity of at least that of TKPPR. Also included are an ultrasound
XX contrast agent (c1) comprising a suspension of gas filled microbubbles
XX comprising the TKPPR compound, an ultrasound contrast agent (c2)
XX comprising a suspension of gas filled microballoons comprising the TKPPR
XX compound, preparation of the TKPPR compound (which comprises conjugating
XX the monomer, multimer or polymer of TKPPR or its analogue with a linker
XX to obtain a compound of formula A-L, forming a covalent or non-covalent
XX bond between A-L and the substrate B 1 or forming a covalent bond between
XX B 1 and the linker to form a conjugate B-L followed by conjugation with

CC the monomer), and a kit for preparing a radiopharmaceutical comprising
CC the compound. The compound used for targeting endothelial cells, tumour
CC cells or other cells which express NP-1, for inhibiting angiogenesis, for
CC ultrasound imaging, staging a tumour, screening at least one targeted
CC ultrasound contrast agent for the ability to target endothelial cells,
CC tumour cells or other cells which express NP-1, for the therapeutic
CC delivery in vivo of a bioactive agent and for delivering desired nucleic
CC acids to endothelial cells, tumour cells or other cells which express NP-
CC 1. The composition is also useful for visualisation therapy or
CC radiotherapy of endothelial cells. The present sequence is a cyclic TKPPR
XX analogue peptide.
XX
XX Sequence 7 AA;
XX
XX Query Match 100.0%; Score 29; DB 7; Length 7;
XX Best Local Similarity 100.0%; Pred. No 1.4e+06;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 TKPPR 5
Db 2 TKPPR 6

RESULT 22
ADD10686

ID ADD10686 standard; peptide; 7 AA.
XX
XX ADD10686;
XX
DT 01-JAN-2004 (first entry)
XX
XX Tuftsin analogue peptide #2.
XX
XX Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
XX antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
XX visualisation therapy; radiotherapy.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Modified-site 1
FT /label= OTHER
FT /note= "Glu is covalently linked to a DPPE
FT (dipalmitoylphosphatidylethanolamine) moiety"
XX
XX US2002147136-A1.
XX
XX 10-OCT-2002.
XX
XX 04-JUN-2001; 2001US-00871974.
XX
XX 02-JUN-2000; 2000US-00585364.
XX
XX (VWRO/) VON WRONSKI M A.
XX (WARI/) MARINELLI E R.
XX (NUNN/) NUNN A D.
XX (PILL/) PILLAI R.
XX (RAMA/) RAMALINGAM K.
XX (TWE/) TWEEDLE M F.
XX (LIND/) LINDER K.
XX (NANJ/) NANJAPPAN P.
XX (RAJU/) RAJU N.
XX
XX Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
XX Tweedle MF, Linder K, Nanjappan P, Raju N;
XX WPI; 2003-800817/75.
XX
XX Composition used in targeting endothelial cells e.g. tumor cells
XX comprises compounds containing monomers, multimers or polymers of L-
XX arginine-L-threonine-L-lysyl-L-prolyl-L-prolyl.
XX
XX Example 21; Page 34; 85pp; English.

XX The invention relates to a composition (A1) comprising compounds
 CC containing monomers, multimers or polymers of TKPPR (ABDI0684).
 CC Composition (A1) comprises a compound of formula A-L-B 1, where A is the
 CC TKPPR peptide, L is a linker moiety (of formula given in the
 CC specification) and B is a substrate (or a phospholipid group,
 CC derivatisable bead attached to a fluorescent or radioactive marker,
 CC bioactive agent, delivery vehicle for genetic material, drug or
 CC therapeutic, or chelating group (preferably N 4, S 4, N 3 S, N 2 S 2 or
 CC NS 3) comprising oxo-PnAO complexed with 99m Tc). The compound
 CC specifically binds to NP-1 (Vascular endothelial growth factor binding
 CC receptor transmembrane glycoprotein) or cells that express NP-1 with
 CC avidity of at least that of TKPPR. Also included are an ultrasound
 CC contrast agent (C1) comprising a suspension of gas filled microbubbles
 CC comprising the TKPPR compound, an ultrasound contrast agent (C2)
 CC comprising a suspension of gas filled microbubbles comprising the TKPPR
 CC compound, preparation of the TKPPR compound (which comprises conjugating
 CC the monomer, multimer or polymer of TKPPR or its analogue with a linker
 CC to obtain a compound of formula A-L, forming a covalent or non-covalent
 CC bond between A-L and the substrate B 1 or forming a covalent bond between
 CC B 1 and the linker to form a conjugate B-L followed by conjugation with
 CC the monomer), and a kit for preparing a radiopharmaceutical comprising
 CC the compound. The compound used for targeting endothelial cells, tumour
 CC cells or other cells which express NP-1, for inhibiting angiogenesis, for
 CC ultrasound imaging, staging a tumour, screening at least one targeted
 CC ultrasound contrast agent for the ability to target endothelial cells,
 CC tumour cells or other cells which express NP-1, for the therapeutic
 CC delivery in vivo of a bioactive agent and for delivering desired nucleic
 CC acids to endothelial cells, tumour cells or other cells which express NP-
 CC 1. The composition is also useful for visualisation therapy or
 CC radiotherapy of endothelial cells. The present sequence is a TKPPR
 CC analogue peptide.

SQ Sequence 7 AA;

Query Match 100.0%; Score 29; DB 7; Length 7;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 DB 3 TKPPR 7

RESULT 23

AAR88740
 ID AAR88740 standard; peptide; 8 AA.

XX AAR88740;

AC AAR88740;

XX 10-APR-1996 (first entry)

DE Tuftsin antagonist peptide-metal chelator conjugate.

XX Peptide-chelator conjugate; metal chelator; diagnostic imaging;
 KW inflammation; radionuclide; tuftsin; analogue; antagonist.

XX Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..5

FT Modified-site 6 /label= tuftsin_antagonist

FT /label= bala

FT /note= "linking group"

FT Modified-site 7

FT /label= OTHER

FT /note= "the side-chain (epsilon) amino group of Lys at
 FT position 7 forms a peptide bond with the C-terminus of
 FT the tetrapeptide N',N'-dimethylglycyl-Ser-Cys(Acm)-Gly-;
 FT this sidechain group consists of a tripeptide metal
 FT chelator and a linking residue (Gly)"

XX

PN WO9522996-A2.

XX 31-AUG-1995.

PD 24-FEB-1995; 95WO-CA000106.

XX 25-FEB-1994; 94US-00202178.

PR (RESO-) RESOLUTION PHARM INC.

XX Goodbody A, Pollak A;

PI WPI; 1995-311386/40.

XX New peptide-chelator conjugate and complex with traceable metal - used to

XX image sites of inflammation in vivo without significant accumulation on

XX the gastrointestinal tract.

XX Claim 19; Page 21; 23pp; English.

XX The present sequence is that of a specifically claimed peptide- chelator

XX conjugate in which a tuftsin antagonist peptide is coupled to a metal

XX chelator, via a linking group. The chelator serves as a labelling site

XX for radionuclide metals such as technetium-99m. The tuftsin antagonist

XX targets the conjugate to macrophages and neutrophils at sites of

XX inflammation without significant accumulation in the gastrointestinal

XX tract (unlike the native tuftsin tetrapeptide). The conjugate is thus

XX useful for diagnostic imaging of inflammation sites, providing an

XX improved target to background ratio

XX Sequence 8 AA;

SQ Query Match 100.0%; Score 29; DB 2; Length 8;

Best Local Similarity 100.0%; Pred. No. 1.4e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

DB 1 TKPPR 5

RESULT 24

AAR76218

ID AAR76218 standard; peptide; 8 AA.

XX AAR76218;

AC AAR76218;

XX 12-JAN-1996 (first entry)

DE peptide-ligand conjugate #1.

XX Peptide-ligand; metal-labelled imaging agent; maleimide; metal atom;

XX technetium; tissue; cell type; organ.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /label= Pic-Ser

FT WO9513832-A1.

XX 26-MAY-1995.

PD 16-NOV-1994; 94WO-CA000637.

XX 16-NOV-1993; 93US-00152680.

XX (RESO-) RESOLUTION PHARM INC.

XX Pollak A, Dunn-Dufault R;

XX WPI; 1995-200205/26.

XX Compens. for generating metal labelled imaging agents - comprising a
 FT solid support, a linking gp. and a ligand cleavable from the linking gp.
 FT by a metal.
 XX
 XX
 PS Claim 6; Page 21; 31pp; English.
 XX
 XX Peptides AAS76218-9 are examples of peptide-ligands used in a method to
 CC produce a metal-labelled imaging agent. The method comprises attaching a
 CC ligand-target molecule e.g. a peptide to a solid surface via a linking
 CC group e.g. maleimide. The ligand can incorporate a metal atom e.g.
 CC technetium, which cleaves the linker-ligand bond thus releasing a
 CC labelled peptide. The labelled peptide can then be used for imaging in
 CC tissues, cell types or organs
 XX
 XX Sequence 8 AA;
 SQ

Query Match 100.0%; Score 29; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e-06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPRP 5
 Db 4 TKPRP 8

RESULT 25
 AAW11058
 ID AAW11058 standard; peptide; 8 AA.
 XX
 AC AAW11058;
 XX
 XX 03-JUN-1997 (first entry)
 DT
 XX Leukocyte-targeted peptide derived radionuclide chelator.
 XX
 DE
 XX Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
 KW diagnostic imaging.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "Sarcosine-Ser"
 FT Modified-site 2 /note= "protected with Acn group"
 FT
 FT
 FT
 FT
 XX WO9603427-A1.
 XX
 XX 08-FEB-1996.
 PD
 XX
 XX 28-APR-1995; 95WO-CA000249.
 PF
 XX 22-JUL-1994; 94US-00279155.
 PR
 XX (RESO-) RESOLUTION PHARM INC.
 PA
 XX Pollak A, Goodbody A;
 PI
 XX WPI; 1996-116994/12.
 DR
 XX New peptide derived radionuclide chelators and metal complexes - useful
 FT for diagnostic imaging.
 PT
 PS Example 3; Page 14; 30pp; English.
 XX
 XX AAW11054-W11059 are peptide derived radionuclide chelators that are
 CC targeted to leukocytes, via the GYKPR sequence. When the chelators are
 CC coupled to a targeting molecule and labelled with a diagnostically useful
 CC metal, they can be used to detect pathological conditions by diagnostic
 CC imaging. The leukocyte-targeted chelators are useful for the rapid
 CC imaging of sites of local inflammation. Radionuclides used include 99mTc,
 CC 64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 166Re, 188Re, 198Au, 199Au, 203Pb, 212Pb

CC and 212Bi. The coupling of a targeting agent and radionuclide using a
 CC chelating agent is an alternative to the direct labelling of targeting
 CC agents in which radionuclides are typically bound at the more numerous
 CC low-affinity sites, forming unstable complexes. The new conjugates give
 CC better scintigraphic images in rat inflammation studies than known
 CC imaging agents Ga-67, 99mTc-19G, 111In-WBC and 99mTc-Nanocol. They image
 CC more rapidly than the known agents and show superior biodistribution
 XX
 SQ Sequence 8 AA;
 Query Match 100.0%; Score 29; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPRP 5
 Db 4 TKPRP 8

RESULT 26
 AAY49844
 ID AAY49844 standard; peptide; 8 AA.
 XX
 AC AAY49844;
 XX
 XX 20-JAN-2000 (first entry)
 DT
 XX Tuftsin receptor antagonist chelate conjugate #5.
 DE
 XX Tuftsin receptor antagonist; chelate conjugate; radiopharmaceutical;
 KW diagnosis; infection; inflammation; imaging; cancer; tumour.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Modified-site 1 /note= "6-(hydrazino)-3-pyridylcarbonyl"
 FT Modified-site 2 /note= "modification on the COOH side group: -Acp-Thr-Lys
 FT Modified-site 3 /label= Acp
 FT /note= "6-aminocaproic acid"
 FT
 FT
 XX WO9951628-A1.
 XX
 XX 14-OCT-1999.
 PD
 XX
 XX 29-MAR-1999; 99WO-US006824.
 PF
 XX 03-APR-1998; 98US-0080672P.
 PR
 XX (DUPO) DU PONT PHARM CO.
 PA
 XX Edwards DS, Rajopadhye M;
 PI
 XX WPI; 1999-633729/54.
 DR
 XX New polypeptides and radiopharmaceuticals used for imaging infection,
 PT inflammation and cancer.
 PT
 XX Claim 4; Page 71; 80pp; English.
 PS
 XX The present invention describes polypeptide compounds of formula (I),
 CC capable of direct transformation into a radiopharmaceutical: Ch-Ln-(X1-X2
 CC -X3-X4-X5)d (I), where X1-X5 = amino acids; Ln = a linking group; Ch = a
 CC metal bonding unit; and d is selected from 1, 2 and 3. The
 CC radiopharmaceuticals are useful for the diagnosis of infection,
 CC inflammation and cancer. The radiopharmaceuticals bind in vivo to the
 CC tuftsin receptor on the surface of white cells which accumulate at the
 CC site of infection and inflammation and can then be detected using
 CC radiation detecting probes or by imaging using a planar or ring gamma
 CC camera. The radiopharmaceuticals can also be used in treating cancer. The

CC present sequence represents a specifically claimed tuftsin receptor
 CC antagonist chelate conjugate from the present invention

XX Sequence 8 AA;

Query Match 100.0%; Score 29; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 4 TKPPR 8

RESULT 27

ABB08448

ID ABB08448 standard; peptide; 8 AA.

XX

AC ABB08448;

XX

XX

DT 01-JUL-2002 (first entry)

XX Tuftsin receptor antagonist (TKPPR) derivative tetramer (BRU-346).

DE

XX Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;

KW angiogenesis; tumour cell; cytostatic; antagonist.

XX

XX

OS Synthetic.

XX

FH Key

FT Modified-site

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

CC related disorder. They may be used for delivering desired nucleic acids
 CC to endothelial cells, tumour cells or other cells expressing NP-1, for
 CC enhancing endothelial or tumour cell-targeted gene therapy, or gene
 CC therapy targeting angiogenic cells, and for treating a human or animal
 CC with a tumour or angiogenesis-related disease. The current sequence
 CC represents a tuftsin receptor antagonist (TKPPR) derivative tetramer BRU-
 CC 346

XX Sequence 8 AA;

Query Match 100.0%; Score 29; DB 5; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||

Db 1 TKPPR 5

RESULT 28

AAR85539

ID AAR85539 standard; peptide; 9 AA.

XX

AC AAR85539;

XX

DT 17-APR-1996 (first entry)

XX

XX

DE Metal chelating peptide conjugated onto tuftsin antagonist.

XX

KW chelate; chelating; chelator; diagnostic; imaging; therapy; tuftsin.

XX

OS Synthetic.

XX

FH Key

FT Domain

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

CC present sequence represents a specifically claimed tuftsin receptor
 CC antagonist chelate conjugate from the present invention

XX Sequence 8 AA;

Query Match 100.0%; Score 29; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 4 TKPPR 8

RESULT 27

ABB08448

ID ABB08448 standard; peptide; 8 AA.

XX

AC ABB08448;

XX

XX

DT 01-JUL-2002 (first entry)

XX Tuftsin receptor antagonist (TKPPR) derivative tetramer (BRU-346).

DE

XX Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;

KW angiogenesis; tumour cell; cytostatic; antagonist.

XX

XX

OS Synthetic.

XX

FH Key

FT Modified-site

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

CC present sequence represents a specifically claimed tuftsin receptor
 CC antagonist chelate conjugate from the present invention

XX Sequence 8 AA;

Query Match 100.0%; Score 29; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 4 TKPPR 8

RESULT 27

ABB08448

ID ABB08448 standard; peptide; 8 AA.

XX

AC ABB08448;

XX

XX

DT 01-JUL-2002 (first entry)

XX Tuftsin receptor antagonist (TKPPR) derivative tetramer (BRU-346).

DE

XX Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;

KW angiogenesis; tumour cell; cytostatic; antagonist.

XX

XX

OS Synthetic.

XX

FH Key

FT Modified-site

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

CC present sequence represents a specifically claimed tuftsin receptor
 CC antagonist chelate conjugate from the present invention

XX Sequence 8 AA;

Query Match 100.0%; Score 29; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 4 TKPPR 8

RESULT 27

ABB08448

ID ABB08448 standard; peptide; 8 AA.

XX

AC ABB08448;

XX

XX

DT 01-JUL-2002 (first entry)

XX Tuftsin receptor antagonist (TKPPR) derivative tetramer (BRU-346).

DE

XX Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;

KW angiogenesis; tumour cell; cytostatic; antagonist.

XX

XX

OS Synthetic.

XX

FH Key

FT Modified-site

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

CC binding peptide. The molecule can be used to chelate a diagnostically or
 CC therapeutically useful metal such as 99m-Tc, and the targeting moiety
 CC can be used to direct the molecule to a site of interest within the body
 CC for in-vivo diagnostic imaging or for therapy. The present sequence
 CC represents specifically claimed examples of the new peptides, in which
 CC the targeting peptide is a tuftsin antagonist

SQ Sequence 9 AA;
 Query Match 100.0%; Score 29; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. NO. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 5 TKPPR 9

RESULT 29
 AAR85535
 ID AAR85535 standard; peptide; 9 AA.

AC AAR85535;
 XX AAR85535;
 XX 17-APR-1996 (first entry)

DE Metal chelating peptide conjugated onto tuftsin antagonist.
 XX chelate; chelating; chelator; diagnostic; imaging; therapy; tuftsin.

XX Synthetic.
 XX Key Location/Qualifiers
 FH Domain 1..4 /label= metal_chelating_moiety

FT Modified-site 1 /label= OTHER
 FT /note= "picolinic acid residue"

FT Modified-site 3 /note= "Cys(Acm)"
 FT Domain 5..9 /label= tuftsin_antagonist

XX WO9517419-A1.
 PN 29-JUN-1995.

XX 22-DEC-1994; 94WO-CA000718.
 XX 22-DEC-1993; 93US-00171737.

XX (RESO-) RESOLUTION PHARM INC.
 PA Pollak A, Goodbody A;

PI WPI; 1995-240607/31.
 DR New peptide analogue metal chelator(s) - used for chelating radioactive
 PT metals for in-vivo diagnostic imaging or for therapy.

PS Claim 11; Page 26; 34pp; English.
 XX The patent discloses new metal-chelating peptide analogues of formula

CC Chel-Xaa-Cys-X, in which: Chel is a chelating moiety which is the residue
 CC of a heterocyclic amino acid analogue such as picolinic acid, dipicolinic
 CC acid, chelidonic acid, 2-carboxypyrazine, 2-carboxypyrimidine, 2-
 CC carboxypyrrrole, 2-quinolinic acid or 2- or 3-isoxanthinonic acid; Xaa is
 CC an amino acid residue; the Cys residue is optionally protected; and X is
 CC OH, alkoxyl, an amino acid residue or a targeting moiety (TM). Preferably
 CC Xaa is Gly or Ser; X is Gly-OH or Gly-TM; and TM is a cell receptor
 CC binding peptide. The molecule can be used to chelate a diagnostically or
 CC therapeutically useful metal such as 99m-Tc, and the targeting moiety
 CC can be used to direct the molecule to a site of interest within the body

CC for in-vivo diagnostic imaging or for therapy. The present sequence
 CC represents a specifically claimed example of the new peptides, in which
 CC the targeting peptide is a tuftsin antagonist

SQ Sequence 9 AA;
 Query Match 100.0%; Score 29; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. NO. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
 |||||
 Db 5 TKPPR 9

RESULT 30
 AAR88735
 ID AAR88735 standard; peptide; 9 AA.

AC AAR88735;
 XX AAR88735;
 XX 10-APR-1996 (first entry)

DE Tuftsin antagonist peptide-metal chelator conjugate.
 XX Peptide-chelator conjugate; metal chelator; diagnostic imaging;
 XX inflammation; radionuclide; tuftsin; analogue; antagonist.

XX Synthetic.
 XX Key Location/Qualifiers
 FH Region 1..3 /label= chelator

FT Misc-difference 1 /note= "pref. Chelates a radionuclide"

FT Modified-site 3 /label= OTHER
 FT /note= "picolinic acid or N',N'-dimethylglycine"

FT Modified-site 3 /label= OTHER
 FT /note= "Cys(Acm)"

FT Region 4 /label= linking_group
 FT Peptide 5..9 /label= tuftsin_antagonist

XX WO9522996-A2.
 PN 31-AUG-1995.

XX 24-FEB-1995; 95WO-CA000106.
 XX 25-FEB-1994; 94US-00202178.

XX (RESO-) RESOLUTION PHARM INC.
 PA Goodbody A, Pollak A;

PI WPI; 1995-311386/40.
 DR New peptide-chelator conjugate and complex with traceable metal - used to
 PT image sites of inflammation in vivo without significant accumulation on
 PT the gastrointestinal tract.

PS Claim 16; Page 21; 23pp; English.
 XX The present sequence is that of a specifically claimed peptide- chelator

CC conjugate in which a tuftsin antagonist peptide is coupled to a metal
 CC chelator, via a linking group. The chelator serves as a labelling site
 CC for radionuclide metals such as technetium-99m. The tuftsin antagonist
 CC targets the conjugate to macrophages and neutrophils at sites of
 CC inflammation without significant accumulation in the gastrointestinal
 CC tract (unlike the native tuftsin tetrapeptide). The conjugate is thus
 CC useful for diagnostic imaging of inflammation sites, providing an

```
CC improved target to background ratio
XX
SQ Sequence 9 AA;
Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. NO. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
Db 5 TKPPR 9
RESULT 31
AAR88738
ID AAR88738 standard; peptide; 9 AA.
XX
AC AAR88738;
XX
DT 10-APR-1996 (first entry)
XX
DE Tuftsin antagonist peptide-metal chelator conjugate.
XX
KW Peptide-chelator conjugate; metal chelator; diagnostic imaging;
inflammation; radionuclide; tuftsin; analogue; antagonist.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Region 1..3 /label= chelator
FT /note= "pref. chelates a radionuclide"
FT Modified-site 1 /label= OTHER
FT /note= "N',N'-dimethylglycine"
FT Modified-site 3 /label= OTHER
FT /note= "Cys(Acm)"
FT Modified-site 4 /label= bala
FT /note= "linking group"
FT Peptide 5..9 /label= tuftsin_antagonist
XX
PN WO9522996-A2.
XX
PD 31-AUG-1995.
XX
PF 24-FEB-1995; 95WO-CA000106.
XX
PR 25-FEB-1994; 94US-00202178.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Goodbody A, Pollak A;
XX
DR WPI; 1995-311386/40.
XX
PT New peptide-chelator conjugate and complex with traceable metal - used to
image sites of inflammation in vivo without significant accumulation on
the gastrointestinal tract.
XX
PS Claim 16; Page 21; 23pp; English.
XX
CC The present sequence is that of a specifically claimed peptide- chelator
conjugate in which a tuftsin antagonist peptide is coupled to a metal
chelator via a linking group. The chelator serves as a labelling site
for radionuclide metals such as technetium-99m. The tuftsin antagonist
targets the conjugate to macrophages and neutrophils at sites of
inflammation without significant accumulation in the gastrointestinal
tract (unlike the native tuftsin tetrapeptide). The conjugate is thus
useful for diagnostic imaging of inflammation sites, providing an
improved target to background ratio
CC
```

```
XX
SQ Sequence 9 AA;
Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. NO. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
Db 5 TKPPR 9
RESULT 32
AAR88741
ID AAR88741 standard; peptide; 9 AA.
XX
AC AAR88741;
XX
DT 10-APR-1996 (first entry)
XX
DE Tuftsin antagonist peptide-metal chelator conjugate.
XX
KW Peptide-chelator conjugate; metal chelator; diagnostic imaging;
inflammation; radionuclide; tuftsin; analogue; antagonist.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..5 /label= tuftsin_antagonist
FT Region 6..7 /label= linking_group
FT Modified-site 6 /label= bala
FT Modified-site 7 /label= bala
FT Modified-site 8 /label= OTHER
FT /note= "the side-chain (epsilon) amino group of Lys at
position 8 forms a peptide bond with the C-terminus of
the tetrapeptide N',N'-dimethylglycyl-Ser-Cys(Acm)-Gly-;
this sidechain group consists of a tripeptide metal
chelator and a linking residue (Gly)"
XX
PN WO9522996-A2.
XX
PD 31-AUG-1995.
XX
PF 24-FEB-1995; 95WO-CA000106.
XX
PR 25-FEB-1994; 94US-00202178.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Goodbody A, Pollak A;
XX
DR WPI; 1995-311386/40.
XX
PT New peptide-chelator conjugate and complex with traceable metal - used to
image sites of inflammation in vivo without significant accumulation on
the gastrointestinal tract.
XX
PS Claim 19; Page 21; 23pp; English.
XX
CC The present sequence is that of a specifically claimed peptide- chelator
conjugate in which a tuftsin antagonist peptide is coupled to a metal
chelator via a linking group. The chelator serves as a labelling site
for radionuclide metals such as technetium-99m. The tuftsin antagonist
targets the conjugate to macrophages and neutrophils at sites of
inflammation without significant accumulation in the gastrointestinal
tract (unlike the native tuftsin tetrapeptide). The conjugate is thus
useful for diagnostic imaging of inflammation sites, providing an
improved target to background ratio
CC
```

```

XX SQ Sequence 9 AA;
Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 1 TKPPR 5

RESULT 33
AAR76219
ID AAR76219 standard; peptide; 9 AA.
XX
AC AAR76219;
XX
DT 12-JAN-1996 (first entry)
XX
DE Peptide-ligand conjugate #2.
XX
KW Peptide-ligand; metal-labelled imaging agent; maleimide; metal atom;
KW technetium; tissue; cell type; organ.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 3 /label= N,N'-dimethyl-Gly
FT Modified-site 3
XX
PN WO9513832-A1.
XX
PD 26-MAY-1995.
XX
PF 16-NOV-1994; 94WO-CA000637.
XX
PR 16-NOV-1993; 93US-00152680.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Pollak A, Dunn-Dufault R;
XX
PS WPI; 1995-200205/26.
XX
PT Compsns. for generating metal labelled imaging agents - comprising a
PT solid support, a linking gp. and a ligand cleavable from the linking gp.
PT by a metal.
XX
PS Claim 6; Page 21; 31pp; English.
XX
CC Peptides AAR76218-9 are examples of peptide-ligands used in a method to
CC produce a metal-labelled imaging agent. The method comprises attaching a
CC ligand-target molecule e.g. a peptide to a solid surface via a linking
CC group e.g. maleimide. The ligand can incorporate a metal atom e.g
CC technetium, which cleaves the linker-ligand bond thus releasing a
CC labelled peptide. The labelled peptide can then be used for imaging in
CC tissues, cell types or organs
XX
SQ Sequence 9 AA;
Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 5 TKPPR 9

RESULT 34
AAW11055
ID AAW11055 standard; peptide; 9 AA.

```

```

XX AAW11055;
XX
DT 03-JUN-1997 (first entry)
XX
DE Leukocyte-targeted peptide derived radionuclide chelator.
XX
KW Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
KW diagnostic imaging.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT Modified-site 3 /note= "N, N-dimethyl-Gly"
FT Modified-site 3
FT Modified-site 3 /note= "protected with Acn group"
XX
PN WO9603427-A1.
XX
PD 08-FEB-1996.
XX
PF 28-APR-1995; 95WO-CA000249.
XX
PR 22-JUL-1994; 94US-00279155.
XX
PA (RESO-) RESOLUTION PHARM INC.
XX
PI Pollak A, Goodbody A;
XX
PS WPI; 1996-116994/12.
XX
PT New peptide derived radionuclide chelators and metal complexes - useful
PT for diagnostic imaging.
XX
PS Claim 28; Page 22; 30pp; English.
XX
CC AAW11054-W11059 are peptide derived radionuclide chelators that are
CC targeted to leukocytes, via the GTPPR sequence. When the chelators are
CC targeted to a targeting molecule and labelled with a diagnostically useful
CC metal, they can be used to detect pathological conditions by diagnostic
CC imaging. The leukocyte-targeted chelators are useful for the rapid
CC imaging of sites of local inflammation. Radionuclides used include 99mTc,
CC 64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 188Re, 188Re, 198Au, 203Pb, 212Pb,
CC and 212Bi. The coupling of a targeting agent and radionuclide using a
CC chelating agent is an alternative to the direct labelling of targeting
CC agents in which radionuclides are typically bound at the more numerous
CC low-affinity sites, forming unstable complexes. The new conjugates give
CC better scintigraphic images in rat inflammation studies than known
CC imaging agents Ga-67, 99mTc-199, 111In-WBC and 99mTc-Nanocol. They image
CC more rapidly than the known agents and show superior biodistribution
XX
SQ Sequence 9 AA;
Query Match 100.0%; Score 29; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5
DB 5 TKPPR 9

RESULT 35
AAW11059
ID AAW11059 standard; peptide; 9 AA.
XX
AC AAW11059;
XX
DT 03-JUN-1997 (first entry)
XX
DE Leukocyte-targeted peptide derived radionuclide chelator.
XX

```

KW Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
 KW diagnostic imaging.
 XX
 OS Synthetic.
 XX
 XX Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site 3 /note= "Sarcosine-Gly"
 FT Modified-site 3 /note= "protected with Acm group"
 FT
 XX WO9603427-A1.
 PN
 XX 08-FEB-1996.
 PD
 XX 28-APR-1995; 95WO-CA000249.
 PF
 XX 22-JUL-1994; 94US-00279155.
 PR
 XX (RESO-) RESOLUTION PHARM INC.
 PA
 XX Pollak A, Goodbody A;
 PI
 XX WPI; 1996-116994/12.
 DR
 XX New peptide derived radionuclide chelators and metal complexes - useful
 PT for diagnostic imaging.
 PT
 XX Example 1; Page 12; 30pp; English.
 PS
 XX AAW11054-W11059 are peptide derived radionuclide chelators that are
 CC targeted to leukocytes, via the GRKPPR sequence. When the chelators are
 CC coupled to a targeting molecule and labelled with a diagnostically useful
 CC metal, they can be used to detect pathological conditions by diagnostic
 CC imaging. The leukocyte-targeted chelators are useful for the rapid
 CC imaging of sites of local inflammation. Radionuclides used include 99mTc,
 CC 64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 186Re, 188Re, 198Au, 203Pb, 212Pb
 CC and 212Bi. The coupling of a targeting agent and radionuclide using a
 CC chelating agent is an alternative to the direct labelling of targeting
 CC agents in which radionuclides are typically bound at the more numerous
 CC low-affinity sites, forming unstable complexes. The new conjugates give
 CC better scintigraphic images in rat inflammation studies than known
 CC imaging agents Ga-67, 99mTc-IgG, 111In-WBC and 99mTc-Nanocol. They image
 CC more rapidly than the known agents and show superior biodistribution
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 29; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 5 TKPPR 9
 RESULT 36
 AAW11054
 ID AAW11054 standard; peptide; 9 AA.
 XX
 AC AAW11054;
 XX
 XX 03-JUN-1997 (first entry)
 DT
 XX Leukocyte-targeted peptide derived radionuclide chelator.
 DE
 XX Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
 KW diagnostic imaging.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site 3 /note= "N, N-diethyl-Gly"
 FT Modified-site 3 /note= "protected with Acm group"
 FT
 XX WO9603427-A1.
 PN
 XX 08-FEB-1996.
 PD

FT Modified-site 3 /note= "N, N-dimethyl-Gly"
 FT Modified-site 3 /note= "protected with Acm group"
 XX
 PN WO9603427-A1.
 XX
 PD 08-FEB-1996.
 XX
 PF 28-APR-1995; 95WO-CA000249.
 XX
 PR 22-JUL-1994; 94US-00279155.
 XX
 PA (RESO-) RESOLUTION PHARM INC.
 XX
 XX Pollak A, Goodbody A;
 PI
 XX WPI; 1996-116994/12.
 DR
 XX New peptide derived radionuclide chelators and metal complexes - useful
 PT for diagnostic imaging.
 PT
 XX Claim 28; Page 22; 30pp; English.
 PS
 XX AAW11054-W11059 are peptide derived radionuclide chelators that are
 CC targeted to leukocytes, via the GRKPPR sequence. When the chelators are
 CC coupled to a targeting molecule and labelled with a diagnostically useful
 CC metal, they can be used to detect pathological conditions by diagnostic
 CC imaging. The leukocyte-targeted chelators are useful for the rapid
 CC imaging of sites of local inflammation. Radionuclides used include 99mTc,
 CC 64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 186Re, 188Re, 198Au, 203Pb, 212Pb
 CC and 212Bi. The coupling of a targeting agent and radionuclide using a
 CC chelating agent is an alternative to the direct labelling of targeting
 CC agents in which radionuclides are typically bound at the more numerous
 CC low-affinity sites, forming unstable complexes. The new conjugates give
 CC better scintigraphic images in rat inflammation studies than known
 CC imaging agents Ga-67, 99mTc-IgG, 111In-WBC and 99mTc-Nanocol. They image
 CC more rapidly than the known agents and show superior biodistribution
 XX
 SQ Sequence 9 AA;
 Query Match 100.0%; Score 29; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB 5 TKPPR 9
 RESULT 37
 AAW11056
 ID AAW11056 standard; peptide; 9 AA.
 XX
 AC AAW11056;
 XX
 XX 03-JUN-1997 (first entry)
 DT
 XX Leukocyte-targeted peptide derived radionuclide chelator.
 DE
 XX Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
 KW diagnostic imaging.
 KW
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FT Modified-site 1
 FT Modified-site 3 /note= "N, N-diethyl-Gly"
 FT Modified-site 3 /note= "protected with Acm group"
 FT
 XX WO9603427-A1.
 PN
 XX 08-FEB-1996.
 PD

```

XX 28-APR-1995; 95WO-CA000249.
XX
XX 22-JUL-1994; 94US-00279155.
XX
XX (RESO-) RESOLUTION PHARM INC.
XX
XX Pollak A, Goodbody A;
XX
XX WPI; 1996-116994/12.
XX
XX New peptide derived radionuclide chelators and metal complexes - useful
XX for diagnostic imaging.
XX
XX Claim 28; Page 22; 30pp; English.
XX
XX AAW11054-W11059 are peptide derived radionuclide chelators that are
XX targeted to leukocytes, via the GTPPR sequence. When the chelators are
XX coupled to a targeting molecule and labelled with a diagnostically useful
XX metal, they can be used to detect pathological conditions by diagnostic
XX imaging. The leukocyte-targeted chelators are useful for the rapid
XX imaging of sites of local inflammation. Radionuclides used include 99mTc,
XX 64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 186Re, 188Re, 198Au, 203Pb, 212Pb
XX and 212Bi. The coupling of a targeting agent and radionuclide using a
XX chelating agent is an alternative to the direct labelling of targeting
XX agents in which radionuclides are typically bound at the more numerous
XX low-affinity sites, forming unstable complexes. The new conjugates give
XX better scintigraphic images in rat inflammation studies than known
XX imaging agents Ga-67, 99mTc-IGG, 111In-WBC and 99mTc-Nanocol. They image
XX more rapidly than the known agents and show superior biodistribution
XX
XX Sequence 9 AA;
XX
XX Query Match 100.0%; Score 29; DB 2; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TKPPR 5
XX Db |||||
XX 5 TKPPR 9
XX
XX RESULT 38
XX AAW11057
XX ID AAW11057 standard; peptide; 9 AA.
XX
XX AC AAW11057;
XX
XX DT 03-JUN-1997 (first entry)
XX
XX DE Leukocyte-targeted peptide derived radionuclide chelator.
XX
XX KW Leukocyte; target; direct; chelator; radionuclide; radiolabel; isotope;
XX diagnostic imaging.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Modified-site 1 /note= "N, N-dibenzyl-Gly"
XX FT Modified-site 3 /note= "protected with Acn group"
XX
XX PN WO9603427-A1.
XX
XX PD 08-FEB-1996.
XX
XX PF 28-APR-1995; 95WO-CA000249.
XX
XX PR 22-JUL-1994; 94US-00279155.
XX
XX (RESO-) RESOLUTION PHARM INC.
XX

```

```

PI Pollak A, Goodbody A;
XX
XX WPI; 1996-116994/12.
XX
XX New peptide derived radionuclide chelators and metal complexes - useful
XX for diagnostic imaging.
XX
XX Claim 28; Page 22; 30pp; English.
XX
XX AAW11054-W11059 are peptide derived radionuclide chelators that are
XX targeted to leukocytes, via the GTPPR sequence. When the chelators are
XX coupled to a targeting molecule and labelled with a diagnostically useful
XX metal, they can be used to detect pathological conditions by diagnostic
XX imaging. The leukocyte-targeted chelators are useful for the rapid
XX imaging of sites of local inflammation. Radionuclides used include 99mTc,
XX 64Cu, 67Cu, 97Ru, 105Rh, 109Pd, 186Re, 188Re, 198Au, 203Pb, 212Pb
XX and 212Bi. The coupling of a targeting agent and radionuclide using a
XX chelating agent is an alternative to the direct labelling of targeting
XX agents in which radionuclides are typically bound at the more numerous
XX low-affinity sites, forming unstable complexes. The new conjugates give
XX better scintigraphic images in rat inflammation studies than known
XX imaging agents Ga-67, 99mTc-IGG, 111In-WBC and 99mTc-Nanocol. They image
XX more rapidly than the known agents and show superior biodistribution
XX
XX Sequence 9 AA;
XX
XX Query Match 100.0%; Score 29; DB 2; Length 9;
XX Best Local Similarity 100.0%; Pred. No. 1.4e+06;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 TKPPR 5
XX Db |||||
XX 5 TKPPR 9
XX
XX RESULT 39
XX AAW03420
XX ID AAW03420 standard; peptide; 9 AA.
XX
XX AC AAW03420;
XX
XX DT 10-OCT-1997 (first entry)
XX
XX DE Peptide useful as diagnostic imaging agent.
XX
XX KW chelator; ligand.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Region 1..3 /label= chelator
XX FT Modified-site 1 /note= "N,N-dimethyl-Gly"
XX FT Modified-site 3 /label= ligand derivatised residue
XX /note= "a galactosyl residue is attached to the thiol
XX group of this Cys via a cleavable maleimido-containing
XX coupling group"
XX FT Region 5..9 /label= targeting_molecule
XX
XX PN WO9638185-A1.
XX
XX PD 05-DEC-1996.
XX
XX PF 16-MAY-1996; 96WO-CA000310.
XX
XX PR 31-MAY-1995; 95US-00454859.
XX
XX (RESO-) RESOLUTION PHARM INC.
XX
XX Pollak A;
PI

```

XX WPI; 1997-042662/04.
 XX
 XX New cpd. for use in diagnostic imaging - consists of imaging agent, which
 PT is chelator for traceable metal, coupled by metal-cleavable bond to
 PT ligand which localises at different site.
 XX
 XX Claim 10-14; Page 16; 26pp; English.
 XX
 XX The patent discloses new peptides useful for diagnostic imaging. The
 CC peptides comprise (1) an imaging agent (a chelator for a traceable metal)
 CC which localises selectively at an in-vivo site of diagnostic interest;
 CC (2) a ligand which localises at a point (in vivo) removed from the site
 CC of diagnostic interest; and (3) a metal-cleavable bond coupling the
 CC chelator of the imaging agent to the ligand, this bond being cleaved upon
 CC formation of a coordination complex of the metal and chelator. Preferably
 CC the chelator is the sequence N,N-dimethyl- Gly-Ser-Cys; the ligand is a
 CC galactosyl residue; the traceable metal is ^{99m}Tc; and a targeting
 CC molecule (especially TKPPR) is attached to the chelator. Upon labelling,
 CC the ligand is cleaved, leaving the labelled imaging agent free to
 CC localise at the site of diagnostic interest unhindered, while the ligand
 CC and any unlabelled imaging agent are sequestered to the removed site. By
 CC sequestering unlabelled imaging agent, the labelled imaging agent does
 CC not compete to occupy the site of interest, resulting in images of
 CC enhanced resolution. The present sequence is the preferred molecule used
 CC for imaging, in which a galactosyl residue is attached to the thiol group
 CC of Cys(3) via a cleavable maleimide coupling group
 XX
 XX Sequence 9 AA;
 SQ

Query Match 100.0%; Score 29; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TKPPR 5
 |||||
 Db 5 TKPPR 9

RESULT 40
 AAY23752
 ID AAY23752 standard; peptide; 9 AA.
 AC AAY23752;
 XX
 XX O9-SEP-1999 (first entry)
 DT
 DE Peptide RP502 used to prepare chirally pure peptides.
 XX
 XX Chirally pure peptide; metal complex; diagnostic imaging; radioimaging.
 XX
 XX Synthetic.
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "DimethylGly"
 FT Modified-site 2 /note= "t-butylGly"
 FT Modified-site 3 /note= "Acm protecting group attached"
 FT
 FT
 XX WO9933863-A1.
 PN
 XX
 XX
 PD 08-JUL-1999.
 XX
 XX 23-DEC-1998; 98WO-CA001201.
 PF
 XX 24-DEC-1997; 97US-00997802.
 PR 30-DEC-1997; 97CA-0226226.
 XX
 XX (RESO-) RESOLUTION PHARM INC.
 PA
 XX Pollak A, Fauconnier T, Wong E;
 PI

XX WPI; 1999-419086/35.
 XX
 XX Chirally pure peptides for preparation of metal complexes, used for
 PT diagnostic imaging.
 XX
 XX Example 1; Page 18; 35pp; English.
 PS
 XX The specification describes chirally pure peptides which are used for
 CC preparation of metal complexes. When converted to complexes with
 CC (radioactive) metals or their oxides or nitrides, the peptides are useful
 CC for diagnostic (radio)imaging (by labeling biomolecules). The present
 CC peptide is used in the course of the invention
 XX
 XX Sequence 9 AA;
 SQ

Query Match 100.0%; Score 29; DB 2; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1.4e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TKPPR 5
 |||||
 Db 5 TKPPR 9

RESULT 41
 AAR88739
 ID AAR88739 standard; peptide; 10 AA.
 AC AAR88739;
 XX
 XX 10-APR-1996 (first entry)
 DT
 XX
 XX Tuftsin antagonist peptide-metal chelator conjugate.
 DE
 XX Peptide-chelator conjugate; metal chelator; diagnostic imaging;
 KW inflammation; radionuclide; tuftsin; analogue; antagonist.
 XX
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Region 1, 3
 FT /label= chelator
 FT /note= "pref. chelates a radionuclide"
 FT Modified-site 1 /label= OTHER
 FT Modified-site 3 /note= "N',N'-dimethylglycine"
 FT /label= OTHER
 FT /note= "Cys(Acm) "
 FT Region 4, 5
 FT /label= linking_group
 FT Modified-site 4 /label= bala
 FT Modified-site 5 /label= bala
 FT Peptide 6, 10
 FT /label= tuftsin_antagonist
 XX
 XX WO9522996-A2.
 PN
 XX
 XX 31-AUG-1995.
 PD
 XX 24-FEB-1995; 95WO-CA000106.
 PF
 XX 25-FEB-1994; 94US-00202178.
 PR (RESO-) RESOLUTION PHARM INC.
 XX
 XX Goodbody A, Pollak A;
 PI
 XX WPI; 1995-311386/40.
 DR
 XX

AA
PN
US2003059422-A1.

Tufts in metalloproteinase, useful e.g. for imaging site of infection or inflammation, comprises metal ion-binding backbone including at least two

PD 27-MAR-2003.
 PF 30-AUG-1999; 99US-00387715.
 XX
 PR 07-JUN-1995; 95US-00476652.
 FR 05-JUN-1996; 96US-00660697.
 PR 18-MAR-1998; 98US-0078373P.
 PR 14-DEC-1998; 98US-0112235P.
 PR 18-MAR-1999; 99WO-US005693.
 XX
 PA (SHAR/) SHARMA S D.
 XX
 FI Sharma SD;
 XX
 DR WPI; 2003-596563/56.
 XX
 XX
 FT Tuftsin metalloproteinase, useful e.g. for imaging site of infection or
 PT inflammation, comprises metal ion-binding backbone including at least two
 FT contiguous amino acids for complexing with metal ion.
 XX
 PS Claim 17; Page 11; 12pp; English.
 XX
 CC The invention describes a peptide or its salt comprising a metal ion-
 CC binding backbone including at least two contiguous amino acids for
 CC complexing with the metal ion. The peptide is specific for the tuftsin
 CC receptor on complexing the metal ion-binding backbone with the metal ion.
 CC The peptide is useful for imaging a site of infection or inflammation;
 CC for causing an immunostimulatory response in mammals; for treating immune
 CC system disorders; in biological, pharmaceutical and radiopharmaceutical
 CC applications; as an analgesic in the treatment of CNS conditions; and in
 CC whole body imaging and radiotherapy. The peptide complexed with the metal
 CC ion is resistant to enzymatic degradation. The affinity of the peptide
 CC for the tuftsin receptor is higher when the metal ion-binding backbone is
 CC complexed with the metal ion than that when the metal ion-binding
 CC backbone is not complexed with the metal ion. The Tc-labeled Thr-D-Lys-
 CC Gly-D-Cys-Arg is the most potent existing tuftsin molecule. The peptide-
 CC metal ion complexes have a higher level of stability, and are less
 CC susceptible to proteolysis than the uncomplexed or existing peptides. The
 CC peptide analogue is not conformationally restricted in the absence of a
 CC metal ion but has high potency and concomitant conformational restriction
 CC on complexation with a metal ion. The metal complexation in the peptide
 CC causes specific regional conformational restrictions in the peptide, so
 CC that the peptide conformation at the metal binding site is
 CC conformationally fixed on metal complexation. The complexation of the
 CC peptide to a metal ion alters the in vivo distribution profile, rate and
 CC mode of clearance from the body, bioavailability and pharmacokinetics in
 CC mammals. The peptide metal ion complex stimulates polymorphonuclear
 CC granulocytes, monocytes and macrophages towards phagocytosis, produces
 CC higher titer antibodies, and can transit the gut-blood barrier without
 CC significant enzymatic or peptidase degradation. This is the amino acid
 CC sequence of a tuftsin metalloproteinase analogue of the invention.
 XX
 SQ Sequence 10 AA;
 Query Match 100.0%; Score 29; DB 7; Length 10;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB |||||
 6 TKPPR 10
 RESULT 44
 AAR88736
 ID AAR88736 standard; peptide; 11 AA.
 XX
 AC AAR88736;
 XX
 DT 10-APR-1996 (first entry)
 XX
 DE Tuftsin antagonist peptide-metal chelator conjugate.
 XX

KW Peptide-chelator conjugate; metal chelator; diagnostic imaging;
 KW inflammation; radionuclide; tuftsin; analogue; antagonist.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 XX Region 1. .3
 FT /label= chelator
 FT /note= "pref. chelates a radionuclide"
 FT Misc-difference 1
 FT /label= OTHER
 FT /note= "picolinic acid or benzoyl-mercaptoproacetic acid"
 FT Modified-site 3
 FT /label= OTHER
 FT /note= "Cys(ACM) a"
 FT Region 4. .6
 FT /label= linking_group
 FT Peptide 7. .11
 FT /label= tuftsin_antagonist
 XX
 XX WO9522996-A2.
 PN
 XX 31-AUG-1995.
 XX
 PF 24-FEB-1995; 95WO-CA000106.
 XX
 PR 25-FEB-1994; 94US-00202178.
 XX
 PA (RESO-) RESOLUTION PHARM INC.
 XX
 XX Goodbody A, Pollak A;
 XX WPI; 1995-311386/40.
 DR
 XX New peptide-chelator conjugate and complex with traceable metal - used to
 PT image sites of inflammation in vivo without significant accumulation on
 PT the gastrointestinal tract.
 XX
 PS Claim 16; Page 21; 23pp; English.
 XX
 CC The present sequence is that of a specifically claimed peptide-chelator
 CC conjugate in which a tuftsin antagonist peptide is coupled to a metal
 CC chelator, via a linking group. The chelator serves as a labelling site
 CC for radionuclide metals such as technetium-99m. The tuftsin antagonist
 CC targets the conjugate to macrophages and neutrophils at sites of
 CC inflammation without significant accumulation in the gastrointestinal
 CC tract (unlike the native tuftsin tetrapeptide). The conjugate is thus
 CC useful for diagnostic imaging of inflammation sites, providing an
 CC improved target to background ratio
 XX
 SQ Sequence 11 AA;
 Query Match 100.0%; Score 29; DB 2; Length 11;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TKPPR 5
 DB |||||
 7 TKPPR 11
 RESULT 45
 ADD10691
 ID ADD10691 standard; peptide; 11 AA.
 XX
 AC ADD10691;
 XX
 DT 01-JAN-2004 (first entry)
 XX
 DE Tuftsin analogue peptide tetramer.
 XX
 DE Phagocytosis; tuftsin; endothelial cell; inflammation; cytostatic;
 KW antiangiogenic; NP-1; ultrasound contrast agent; tumour; angiogenesis;
 XX

visualisation therapy; radiotherapy.

Synthetic.

Key Modified-site Location/Qualifiers

6
/label= OTHER
/notes "Lys is covalently linked to a further TKPPRKPPT peptide whose Lys (6) is linked to Gly-TTDA (4,7,10-Trioxal,13-Tridecanediamine)-oregon green"

US2002147136-A1.

10-OCT-2002.

04-JUN-2001; 2001US-00871974.

02-JUN-2000; 2000US-00585364.

(VWRO// VON WRONSKI M A.
(MARI// MARINELLI E R.
(NUNN// NUNN A D.
(PILL// PILLAI R.
(RAMA// RAMALINGAM K.
(TWEED// TWEEDLE M F.
(LIND// LINDER K.
(NANJ// NANJAPPAN P.
(RAJU// RAJU N.

Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
Tweedle MF, Linder K, Nanjappan P, Raju N;
WPI; 2003-800817/75.

Composition used in targeting endothelial cells e.g. tumor cells comprises compounds containing monomers, multimers or polymers of L-arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.

Example 29; Page 66; 85pp; English.

The invention relates to a composition (A1) comprising compounds containing monomers, multimers or polymers of TKPPR (ADD10684). Composition (A1) comprises a compound of formula A-L-B 1, where A is the TKPPR peptide, L is a linker moiety (of formula given in the specification), and B is a substrate (or a phospholipid group, derivatisable bead attached to a fluorescent or radioactive marker, bioactive agent, delivery vehicle for genetic material, drug or therapeutic, or chelating group (preferably N 4, S 4, N 3 S, N 2 S 2 or NS 3) comprising oxa-PnAO complexed with 99m Tc). The compound specifically binds to NP-1 (Vascular endothelial growth factor binding receptor transmembrane glycoprotein) or cells that express NP-1 with avidity of at least that of TKPPR. Also included are an ultrasound contrast agent (C1) comprising a suspension of gas filled microbubbles comprising the TKPPR compound, an ultrasound contrast agent (C2) comprising a suspension of gas filled microballoons comprising the TKPPR compound, preparation of the TKPPR compound (which comprises conjugating the monomer, multimer or polymer of TKPPR or its analogue with a linker to obtain a compound of formula A-L, forming a covalent or non-covalent bond between A-L and the substrate B 1 or forming a covalent bond between B 1 and the linker to form a conjugate B-L followed by conjugation with the monomer), and a kit for preparing a radiopharmaceutical comprising the compound. The compound used for targeting endothelial cells, tumour cells or other cells which express NP-1, for inhibiting angiogenesis, for ultrasound imaging, staging a tumour, screening at least one targeted ultrasound contrast agent for the ability to target endothelial cells, tumour cells or other cells which express NP-1 for the therapeutic delivery in vivo of a bioactive agent and for delivering desired nucleic acids to endothelial cells, tumour cells or other cells which express NP-1. The composition is also useful for visualisation therapy or radiotherapy of endothelial cells. The present sequence is a TKPPR tetramer peptide.

Sequence 11 AA;

Query Match 100.0%; Score 29; DB 7; Length 11;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TKPPR 5
|
|
|
|
Db 1 TKPPR 5

RESULT 46
ADD10690
ID ADD10690 standard; peptide; 11 AA.
XX
XX AC ADD10690;
XX
XX DT 01-JAN-2004 (first entry)
XX
XX DE Tuftsin analogue peptide dimer.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
XX FT Modified-site 6
XX FT /label= OTHER
XX FT /notes "Lys is covalently linked to Gly-TTDA (4,7,10-Trioxal,13-Tridecanediamine)"
XX
XX FN US2002147136-A1.
XX
XX PD 10-OCT-2002.
XX
XX PF 04-JUN-2001; 2001US-00871974.
XX
XX PR 02-JUN-2000; 2000US-00585364.
XX
XX PA (VWRO// VON WRONSKI M A.
XX PA (MARI// MARINELLI E R.
XX PA (NUNN// NUNN A D.
XX PA (PILL// PILLAI R.
XX PA (RAMA// RAMALINGAM K.
XX PA (TWEED// TWEEDLE M F.
XX PA (LIND// LINDER K.
XX PA (NANJ// NANJAPPAN P.
XX PA (RAJU// RAJU N.
XX
XX PI Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
XX PI Tweedle MF, Linder K, Nanjappan P, Raju N;
XX
XX DR WPI; 2003-800817/75.
XX
XX PT Composition used in targeting endothelial cells e.g. tumor cells
XX PT comprises compounds containing monomers, multimers or polymers of L-
XX PT arginine-L-threonyl-L-lysyl-L-prolyl-L-prolyl.
XX
XX PS Example 29; Page 66; 85pp; English.
XX
XX CC The invention relates to a composition (A1) comprising compounds
XX CC containing monomers, multimers or polymers of TKPPR (ADD10684).
XX CC Composition (A1) comprises a compound of formula A-L-B 1, where A is the
XX CC TKPPR peptide, L is a linker moiety (of formula given in the
XX CC specification), and B is a substrate (or a phospholipid group,
XX CC derivatisable bead attached to a fluorescent or radioactive marker,
XX CC bioactive agent, delivery vehicle for genetic material, drug or
XX CC therapeutic, or chelating group (preferably N 4, S 4, N 3 S, N 2 S 2 or
XX CC NS 3) comprising oxa-PnAO complexed with 99m Tc). The compound
XX CC specifically binds to NP-1 (Vascular endothelial growth factor binding
XX CC receptor transmembrane glycoprotein) or cells that express NP-1 with
XX CC avidity of at least that of TKPPR. Also included are an ultrasound
XX CC contrast agent (C1) comprising a suspension of gas filled microbubbles
XX CC comprising the TKPPR compound, an ultrasound contrast agent (C2)
XX CC comprising a suspension of gas filled microballoons comprising the TKPPR
XX CC compound, preparation of the TKPPR compound (which comprises conjugating
XX CC the monomer, multimer or polymer of TKPPR or its analogue with a linker
XX CC to obtain a compound of formula A-L, forming a covalent or non-covalent
XX CC bond between A-L and the substrate B 1 or forming a covalent bond between
XX CC B 1 and the linker to form a conjugate B-L followed by conjugation with
XX CC the monomer), and a kit for preparing a radiopharmaceutical comprising
XX CC the compound. The compound used for targeting endothelial cells, tumour
XX CC cells or other cells which express NP-1, for inhibiting angiogenesis, for
XX CC ultrasound imaging, staging a tumour, screening at least one targeted
XX CC ultrasound contrast agent for the ability to target endothelial cells,
XX CC tumour cells or other cells which express NP-1 for the therapeutic
XX CC delivery in vivo of a bioactive agent and for delivering desired nucleic
XX CC acids to endothelial cells, tumour cells or other cells which express NP-
XX CC 1. The composition is also useful for visualisation therapy or
XX CC radiotherapy of endothelial cells. The present sequence is a TKPPR
XX CC tetramer peptide.
XX
XX SQ Sequence 11 AA;

CC contrast agent (c1) comprising a suspension of gas filled microbubbles
 CC comprising the TKPPR compound, an ultrasound contrast agent (c2)
 CC comprising a suspension of gas filled microballoons comprising the TKPPR
 CC compound, preparation of the TKPPR compound (which comprises conjugating
 CC the monomer, multimer or polymer of TKPPR or its analogue with a linker
 CC to obtain a compound of formula A-L, forming a covalent or non-covalent
 CC bond between A-L and the substrate B 1 or forming a covalent bond between
 CC B 1 and the linker to form a conjugate B-L followed by conjugation with
 CC the monomer), and a kit for preparing a radiopharmaceutical comprising
 CC the compound, and a kit for preparing a radiopharmaceutical comprising
 CC cells or other cells which express NP-1, for inhibiting angiogenesis, for
 CC ultrasound imaging, staging a tumour, screening at least one targeted
 CC ultrasound contrast agent for the ability to target endothelial cells,
 CC tumour cells or other cells which express NP-1, for the therapeutic
 CC delivery in vivo of a bioactive agent and for delivering desired nucleic
 CC acids to endothelial cells, tumour cells or other cells which express NP-
 CC 1. The composition is also useful for visualisation therapy or
 CC radiotherapy of endothelial cells. The present sequence is a TKPPR dimer
 CC peptide.
 XX
 XX

SQ Sequence 11 AA;

Query Match 100.0%; Score 29; DB 7; Length 11;
 Best Local Similarity 100.0%; Pred. No. 79;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

DB 1 TKPPR 5

RESULT 47

ABB08446
 ID ABB08446 standard; peptide; 12 AA.

XX ABB08446;

DT 01-JUL-2002 (first entry)

DE Tuftsin receptor antagonist (TKPPR) derivative peptide 3.

XX Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
 KW angiogenesis; tumour cell; cytostatic; antagonist.
 XX

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "residue modified by the following; F-108,
 FT OCH2CONH"

FT Modified-site 2 /note= "residue modified by (tBu)"

FT Modified-site 3 /note= "residue modified by (Mtt)"

FT Modified-site 6 /note= "residue modified by (Mtt)"

FT Modified-site 7 /note= "residue modified by (Pmc)-OtBu"

FT Modified-site 8 /note= "residue modified by the following; F-108,
 FT OCH2CONH"

FT Modified-site 9 /note= "residue modified by (tBu)"

FT Modified-site 12 /note= "residue modified by (Mtt)"

FT Modified-site 12 /note= "residue modified by (Pmc)-OtBu"

XX WO200191805-A2.

XX 06-DEC-2001.

XX 04-JUN-2001; 2001WO-US018053.

XX 02-JUN-2000; 2000US-00585364.

XX (BRAC) BRACCO RES USA.

XX (BRAC) BRACCO RES USA.
 PA Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
 XX Tweedle MF, Linder K, Nanjappan P, Raju N;
 PI WPI; 2002-195523/25.
 XX

DR Composition for use in targeting endothelial cells, tumor cells or other
 PT cells which express NP-1 comprises a compound containing a polypeptide,
 PT linker and substrate.
 XX

PS Example 25; Page 95; 146pp; English.

XX The invention relates to a composition for use in targeting endothelial
 CC cells, tumour cells, or other cells which express NP-1. The activity of
 CC compositions of the invention may be described as cytostatic. Compounds
 CC of the invention are useful in pharmaceutical compositions for inhibiting
 CC angiogenesis, for imaging and targeting an angiogenic site, endothelial
 CC cells, tumour cells or other cells that express NP-1 in a human or
 CC animal. They may also be used as ultrasound contrast agents, for staging
 CC a tumour in a human or animal, for screening for the ability of an agent
 CC to target endothelial cells, tumour cells or other cells that express NP-
 CC 1. They may be used for therapeutic delivery in vivo of a bioactive agent
 CC or for treating an individual exhibiting effects of an angiogenesis or a
 CC related disorder. They may be used for delivering desired nucleic acids
 CC to endothelial cells, tumour cells or other cells expressing NP-1, for
 CC enhancing endothelial or tumour cell-targeted gene therapy, or gene
 CC therapy targeting angiogenic cells, and for treating a human or animal
 CC with a tumour or angiogenesis-related disease. The current sequence
 CC represents a tuftsin receptor antagonist (TKPPR) derivative of the
 CC invention
 XX

SQ Sequence 12 AA;

Query Match 100.0%; Score 29; DB 5; Length 12;
 Best Local Similarity 100.0%; Pred. No. 85;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TKPPR 5

DB 2 TKPPR 6

RESULT 48

ABB08449
 ID ABB08449 standard; peptide; 20 AA.

XX ABB08449;

DT 01-JUL-2002 (first entry)

DE 99mTc radiopharmaceutical.

XX Tuftsin; endothelial cell; drug delivery; gene therapy; NP-1;
 KW angiogenesis; tumour cell; cytostatic; antagonist; radiopharmaceutical.
 XX

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "residue modified by the addition of 99mTc-Oxa
 FT FNAO"

FT WO200191805-A2.

XX 06-DEC-2001.

XX 04-JUN-2001; 2001WO-US018053.

XX 02-JUN-2000; 2000US-00585364.

XX (BRAC) BRACCO RES USA.

XX Von Wronski MA, Marinelli ER, Nunn AD, Pillai R, Ramalingam K;
 PI Tweedle MF, Linder K, Nanjappan P, Raju N;
 XX
 XX WPI; 2002-195523/25.
 XX
 PT Composition for use in targeting endothelial cells, tumor cells or other
 PT cells which express NP-1 comprises a compound containing a polypeptide,
 PT linker and substrate.
 XX
 XX Example 37; Page 128; 146pp; English.
 XX
 CC The invention relates to a composition for use in targeting endothelial
 CC cells, tumor cells, or other cells which express NP-1. The activity of
 CC compositions of the invention may be described as cytostatic. Compounds
 CC of the invention are useful in pharmaceutical compositions for inhibiting
 CC angiogenesis, for imaging and targeting an angiogenic site, endothelial
 CC cells, tumor cells or other cells that express NP-1 in a human or
 CC animal. They may also be used as ultrasound contrast agents, for staging
 CC a tumour in a human or animal, for screening for the ability of an agent
 CC to target endothelial cells, tumour cells or other cells that express NP-
 CC 1. They may be used for therapeutic delivery in vivo of a bioactive agent
 CC or for treating an individual exhibiting effects of an angiogenesis or a
 CC related disorder. They may be used for delivering desired nucleic acids
 CC to endothelial cells, tumour cells or other cells expressing NP-1, for
 CC enhancing endothelial or tumour cell-targeted gene therapy, or gene
 CC therapy targeting angiogenic cells, and for treating a human or animal
 CC with a tumour or angiogenesis-related disease. The current sequence
 CC represents a tuftsin receptor antagonist (TKPPR)-metal chelator conjugate
 CC complexed with a radioactive metal that is described in the specification
 CC as 99mTc radiopharmaceutical
 XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 29; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02; Indels 0; Gaps 0;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TKPPR 5
 Db 1 TKPPR 5
 RESULT 49
 ABP34343
 ID ABP34343 standard; protein; 53 AA.
 AC ABP34343;
 XX
 XX 08-JUL-2002 (first entry)
 DE Human ORF3316 protein, SEQ ID NO:6632.
 XX
 KW Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibitor; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnary;
 KW vasotropic; antipsoriatic; antidiabetic; cytostatic; neurotropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antithyroid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 XX
 PN WC200190366-A2.
 XX
 XX 29-NOV-2001.
 PD
 XX

PF 24-MAY-2001; 2001WO-US017076.
 XX
 PR 24-MAY-2000; 2000US-0206690P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Leach MD, Shinkets RA;
 XX
 XX WPI; 2002-106200/14.
 DR N-PSDB; ABN78369.
 DR
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation.
 XX
 PS Claim 10; Page 1905; 2508pp; English.
 XX
 CC Sequences ABP3028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibitor activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, tumour inhibition activity,
 CC receptor/ligand, antiinflammatory activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases
 XX
 SQ Sequence 53 AA;
 Query Match 100.0%; Score 29; DB 5; Length 53;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 TKPPR 5
 Db 43 TKPPR 47
 RESULT 50
 ABG07007
 ID ABG07007 standard; protein; 65 AA.
 XX
 XX ABG07007;
 AC
 XX
 XX 13-FEB-2002 (first entry)
 DT
 XX Novel human diagnostic protein #6998.
 DE

```

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
FN
XX 11-OCT-2001.
PD
XX 30-YAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR
XX 23-AUG-2000; 2000US-00649167.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI: 2001-639362/73.
DR
XX N-PSDB; AAS711194.
DR
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
PT
XX Claim 20; SEQ ID NO 37366; 103pp; English.
PS
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 65 AA;
Query Match 100.0%; Score 29; DB 4; Length 65;
Best Local Similarity 100.0%; Pred. NO. 3.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TKPPR 5
DB 27 TKPPR 31
Search completed: March 3, 2004, 12:17:18
Job time : 60 secs

```